HOXA5, A CELL FATE TRANSITION REGULATOR, IMPEDES TUMOR INITIATION AND PROGRESSION IN BREAST CANCER

by

Wei Wen Teo

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of

Doctor of Philosophy

Baltimore, Maryland

November, 2015

Abstract

This thesis focuses on understanding the biological functions of *HOXA5* in mammary epithelial differentiation and breast cancer progression. Members of HOX cluster have long been associated with cellular differentiation. Nonetheless, how HOX genes drive differentiation in mammalian cells is poorly understood. The expression of HOXA5 is frequently deregulated in breast cancer, while higher HOXA5 expression predicts better relapse free survival in breast cancer patients. My research has demonstrated that depleting HOXA5, a key regulator during embryogenesis, increases cell plasticity and stemness in the human breast epithelial cell line, MCF10A. Comparison of MCF10A cells and HOXA5-depleted MCF10A cells by expression array analysis revealed HOXA5's ability to regulate several traits of epithelial lineage, including E-cadherin and CD24. Analysis of the expression array by gene set enrichment analysis (GSEA) confirmed the negative correlation between epithelial traits and the transcriptome of HOXA5-depleted MCF10A cells. Using an inducible HOXA5-knockdown system, we demonstrated how HOXA5 depletion restrained state transition from $CD24^-/CD44^+$ to $CD24^+/CD44^+$ cells upon treatment of retinal. The increase of expression of adhesions proteins was, however, impeded when HOXA5 was simultaneously repressed. Failure of state transition upon depletion of HOXA5 resulted in increased ability for mammosphere formation. Furthermore, depleting HOXA5 in MCF10A-KRAS transformed cells reduced $CD24^+/CD44^{lo}$ population and enhanced the

self-renewal capacity of MCF10A-KRAS cells in culture. In Matrigel, these cells exhibited stellate and protrusive morphology which suggested the transition of the cancer cell to a more progenitor-like state. HOXA5-depleted MCF10-KRAS cells failed to develop pseudo-luminal architecture in orthotopically engrafted xenografts. Ectopic expression of HOXA5 in SUM149, on the other hand, resulted in an increased $CD24^{+}/CD44^{-}$ population and inhibition of mammosphere formation in culture and tumor initiation in vivo. Mechanistically, we showed that wild type HOXA5 induced luciferase activity regulated by the proximal promoter region of E-cadherin and CD24 in 293T cells. We further confirmed the recruitment of HOXA5 to putative consensus HOXA5 binding site, TAAT, to these promoters by ChIP analysis. Collectively, these findings support a role for HOXA5 in cell differentiation by up-regulating epithelial traits and suppressing tumor aggressiveness in breast cancer.

Thesis Readers:

Dr. Saraswati Sukumar

Dr. Andrew J. Ewald

Preface & Acknowledgement

I worked as a research specialist for two years from 2007-2009 with Dr. Sara Sukumar before I left to pursue my PhD in the Pathobiology Graduate Program of the Johns Hopkins University School of Medicine. I received an enormous amount of support, both emotional and intellectual, from my mentor and lab colleagues during my training in her laboratory. This became an important factor when in my decision to pursue my PhD studies in the Sukumar laboratory at the Johns Hopkins Breast Cancer Program. I would like to express my gratitude to Dr. Sara Sukumar, my mentor for the past eight years, for giving me an opportunity to pursue my research career at Johns Hopkins University. Her guidance and faith in me were central in allowing me lead the project in studying the biological foundations of HOXA5 action, which is presented in this dissertation. I would also acknowledge Dr. Mary Jo Fackler and Dr. LiangFeng Han for playing a significant role in my development as an independent and versatile scientist. Our laboratory has succeeded in creating a sense of belonging in everyone who has worked here in the past, regardless of the length of their tenure. The friendly environment made the learning process more enjoyable, especially at times when I struggled with my experiments. I feel very thankful to Dr. Vanessa Merino, Dr. Preethi Korangath and Mr. Sean Cho because I would have not culminated my thesis project without generous support from them. My gratitude also goes to past and present lab members including Dr. Kideok Jin, Dr. Xiaohui Liang, Dr. Takahiro Yoshida, Dr. Xinyan Wu, Dr. Helen Sadik, Dr. Leigh Ann Cruz, Ms. Sunju Park, Ms. Sidra Hafeez,

Ms. Zoila Lopez Bujanda, Mr. Hardik Panday, Ms. Veena Padmanabhan, Mr. Benjamin Collins and many others who I could not list herein.

I feel deeply grateful for receiving countless constructive advice and feedback from past and present thesis committee members. Dr. Christine Iacobuzio-Donahue and Dr. Steven Leach offered me valuable comments in the past that helped in shaping my thesis and research career. I am thankful to Dr. Edward Gabrielson and Dr. Akhilesh Pandey for stepping in as the new thesis committee members due to the departure of the previous advisors from Johns Hopkins. Their recommendations and critique benefited my research progress. The project could not have been completed without the collaboration with my thesis committee member, Dr. Andrew Ewald. His expertise in three dimensional models has been instrumental in driving this project forward. I wish to express my appreciation to Dr. Ewald and his graduate student, Mr. Neil Neumann, for their support on studies involving confocal imaging.

I would have not begun my doctoral training at Johns Hopkins University if not for the financial support I received from Mr. Al Njoo and his wife Madam Margaret Lee. I would like to express my thankfulness to them for their generosity and commitment in funding scholars in the Pathobiology Graduate Program. It is an honor to pursue my PhD degree at Pathobiology Graduate Program and I am deeply grateful to our Program Director, Dr. Noel Rose; the Pathology Department Director, Dr. Ralph Hruban; my program advisor, Dr. Kathleen Gabrielson; and our program coordinators Ms. Wihelmena Braswell, Ms. Nancy Nath, and Ms. Tracie McElroy.

I am thankful for the immense help from friends and classmates during the journey pursuing my PhD. Much of my gratitude goes to my classmate, Dr. Ren-Chin

Wu, from whom I have received a tremendous amount of insightful advice on many aspects. My family has always been very supportive of my career choice, and I am blessed to have open-minded parents who have patiently showered me with a lot of love while I was working towards my graduation. I also want to express my thankfulness to my wife, Kate, whom I met serendipitously at Johns Hopkins SOM in 2013. With her continued love and confidence in me, I am able to bring down the curtain on my 6 years of graduate study.

Thank you.

Thesis Advisor:

Dr. Saraswati Sukumar

Thesis Committee Members:

Dr. Andrew J. Ewald

Dr. Akhilesh Pandey

Dr. Edward Gabrielson

Table of Contents

List of Figures

1 Introduction

HOX genes belong to the largest homeodomain containing superfamily whose members are primarily responsible for establishing cell fate in multicellular organism. Members of the HOX clusters were first identified during 1970-1980 in Drosophila through a series of experiments conducted by Edward B. Lewis, Christiane Nüsslein-Volhard and Eric F. Wieschaus. Mutants of HOX genes transform the body patterning of Drosophila during embryogenesis, and hence they were given the name-homeotic

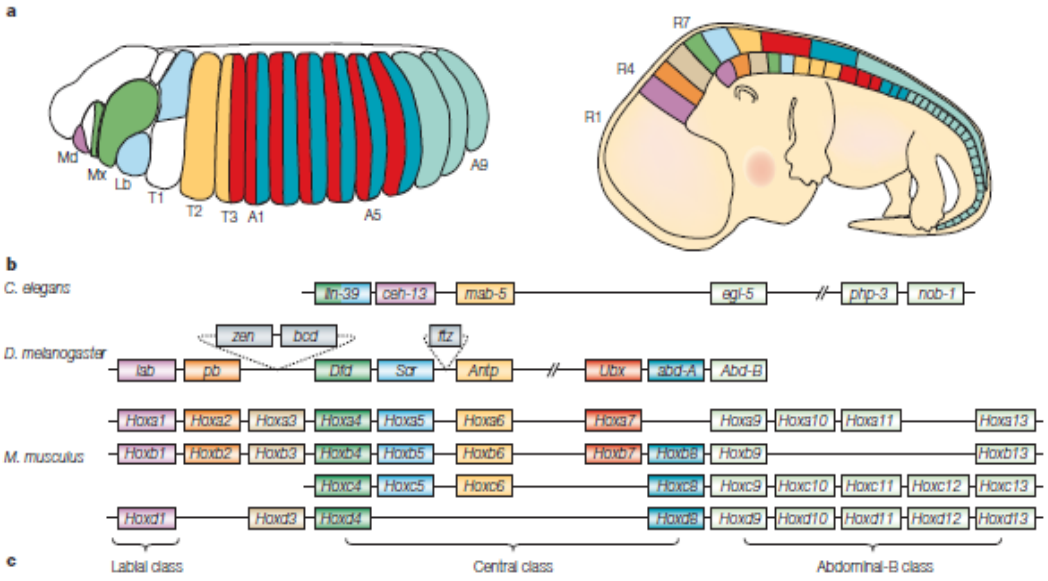


Figure 1. *HOX gene clusters*

genes (maintain homeosis). Unlike the other homeotic genes, HOX genes are clustered in proximity in the genome. There are four HOX (A-D) clusters in human genome, which are located at chromosome 2, 12, 7 and 17 respectively (Figure 1). It is generally believed that these clusters evolved from a common ancestor during evolution. It is also fascinating that the order of expression of the HOX genes follows the sequence of their

positions on the chromosome in the developing embryo. But the purpose for this collinearity and the mechanism of this regulation are still poorly understood. Retinoic acid is a potent differentiating factor which is known to control the collinear expression of HOX genes in teratoma cell line (Simeone et al 1990, Stornaiuolo et al 1990). The connection between retinoids and HOX genes suggests that the latter play an intermediary role in regulating differentiation in mammalian cells. However, the role of HOX genes in the differentiating cascades induced by retinoic acid remains poorly understood.

Members of both HOX and retinoic acid receptor (RAR) families are known to be associated with cellular differentiation (Kondo et al 1992, Marshall et al 1996, Pratt et al 1993, Zwartkruis et al 1993). HOX genes are transcription factors that orchestrate body segmentation during embryogenesis (Alexander et al 2009, Dubrulle et al 2001) and perturbed HOX expression has been implicated in cancer (Shah and Sukumar 2010). The mammary gland is an organ that continues to go through multiple cycles of development during life. Moreover, persistent expression of several HOX members in adulthood suggests the necessity of their biological functions in maintaining tissue homeostasis and differentiation (Taylor et al 1997, Yan et al 2013). Hence, suppression of specific HOX genes in adulthood could be detrimental to normal mammary gland development.

Sex comb reduced (Scr), the ortholog of HOXA5 in Drosophila, promotes normal development of the salivary gland (Haberman et al 2003, Henderson and Andrew 2000). Scr is primarily expressed in the primitive cells that form the salivary gland. However, Scr mRNA expression ceases as the glandular cells invaginate. In fact,

de novo salivary glands form at new body segments when Scr is ectopically expressed throughout the embryo. It is unclear how the spatial temporal expression Scr regulates the development of salivary gland in Drosophila. On the other hand, Hoxa5-/- mice display severe postnatal abnormalities. Mammary gland development is impaired in the surviving mutant mice, which is the result of the loss of Hoxa5 functions in the stromal compartment (Garin et al 2006).

Among the HOX genes clusters, members of the HOXA cluster are frequently silenced through epigenetic regulation in different types of cancer (Novak et al 2006). Silencing of HOXA5 in breast cancer has been associated with promoter hyper-methylation (Watson et al 2004). Investigations from our laboratory have reported loss of HOXA5 in breast cancer and its tumor suppressive function in promoting apoptosis in breast cancer in a p53-dependent and independent manner (Chen et al 2004, Raman et al 2000). Given the strong influence of retinoic acid in the direct transcriptional activation of HOXA5 (Chen et al 2007) and mammary gland development (Chute et al 2006), we sought to understand the function of HOXA5 in mammary epithelial cell differentiation and its tumor suppressive effects during breast cancer progression.

The majority of apocrine glands including the mammary gland are composed primarily of epithelial cells. Epithelium is characterized by its tightly packed pattern, displaying anterior posterior polarity, by which they serve as a barrier between the outside environment and the body. The epithelial cells are attached to each other through cell-cell junctions. E-cadherin is the major component of adhesion junctions which link intracellular actin filaments via catenins. Another cell-cell junction called tight junctions consists of Occludin and Claudins; the complex anchors intracellular

cytoskeleton through ZO-1. Loss of junction proteins disrupts the polarity of epithelial cells and hence impairs their normal functions. Decrease in cell-cell junctions is a typical hallmark of cancer progression, promotes migration and is likely to promote local and distant metastasis (Jeanes et al 2008, Martin and Jiang 2009).

It has been speculated that a population of cells within a tumor or cancer is capable of initiating tumor fomration independantly. These cells are self-renewal and are capable to generate more differented cells, resembling the characterics of normal stem cells. Research in the last twenty years has identified multiple markers common to both normal and breast cancer stem/progenitor cells. The most commonly used among them is the distribution on the cells of two cell surface markers, $CD24^{-/low}/CD44^{+}$ (Figure 2). The surface markers were discovered through xenograft assays using specific populations isolated from primary breast cancer (Al-Hajj et al 2003). It has been further shown in a number of studies that this subpopulation, which is characterized by high levels of expression of CD44 and low or undetectable levels of CD24, displays robust self-renewal and tumorigenic properties (Dontu et al 2003, Ponti et al 2005). Cells that display $CD24^{-/low}/CD44^{+}$ progenitor cell surface markers exhibit mesenchymal or basal cell properties

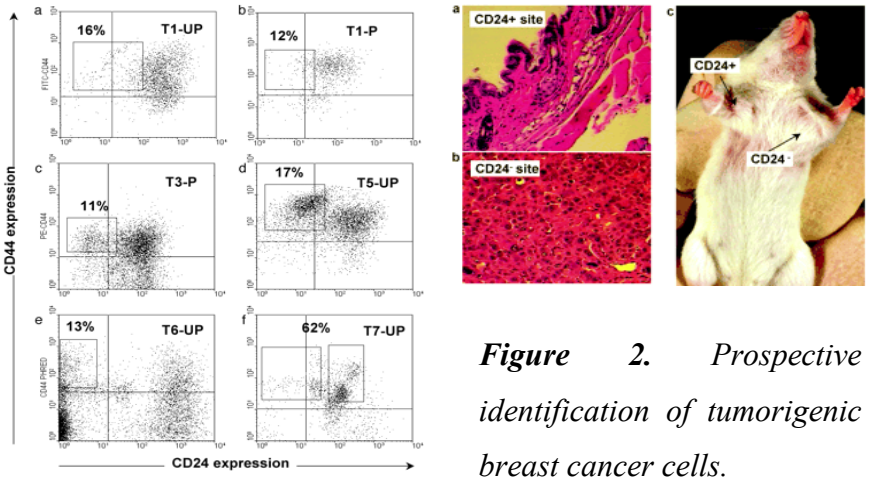


Figure 2. *Prospective identification of tumorigenic breast cancer cells.*

(Mani et al 2008, May et al 2011, Radisky and LaBarge 2008), and are highly invasive and motile compared to their counterparts that express high levels of CD24 (Sheridan et al 2006), express low levels of the adhesion molecule, E-cadherin (Klopp et al 2010, Radisky and LaBarge 2008), and exhibit a cytokeratin profile resembling basal epithelium (Honeth et al 2008).

Among known markers of differentiation, CD24 is glycosyl-phosphatidyl-inositol linked cell surface molecule that has been implicated in lineage development of hematopoietic cells (Kay et al 1991). CD24 expression is low in progenitor/basal cells (Petkova et al 2013, Sleeman et al 2006), but whether it has any biological functions in regulating cell differentiation and specification is as yet unknown. Also, little is known about transcriptional regulation of CD24. The second is retinoic acid, which is a signaling molecule that is produced in the stem cell niche and implicated in stem cell differentiation and induction of lineage progression in hematopoietic cells in vitro (Chute et al 2006, Ghiaur et al 2013). But the pathways utilized by either to mediate differentiation are still not clear.

In this study, we investigated if HOXA5 in breast epithelial cells promotes transitioning of progenitor cells to a differentiated state. Our studies identified loss of expression of several typical epithelial specific genes and loss of the differentiation marker, CD24, occurred upon loss of HOXA5. We provide evidence that HOXA5 acts as a tumor suppressor gene through its ability to regulate cell adhesion through direct transcriptional upregulation of E-cadherin, and promotes lineage determination of stem cells through upregulation of CD24, two major characteristics typical of differentiated mammary epithelial cells.

2 Methods

2.1 Cell lines and Reagents

293T and MCF10A cells were obtained from the American Type Culture Collection, MCF10A-KRAS cells were generated by stably expressing mutant $KRAS^{G12V}$ in MCF10A. SUM149 and SUM159 cells were obtained from Dr. S. Ethier (Wayne State University, Detroit, MI) and a lung metastatic subline of MDA-MB-231 cells called 231-LM2 were obtained Dr. Joan Massagué. MCF10A was maintained in DMEM/F12 supplemented with 5% horse serum, 20 ng/ml epidermal growth factor (#E9644), 500 μg/ml hydrocortisone (#H0888), 10 μg/ml human recombinant insulin (#91077C) and 100 ng/ml cholera toxin (#C8052) from Sigma. SUM149 and SUM159 were maintained in DMEM/F12 supplemented with 5% heat inactivated fetal bovine serum, 10 μg/ml insulin and 500 μg/ml hydrocortisone. 293T and 231-LM2 cells were maintained in DMEM/10% fetal bovine serum. All cells were maintained in medium containing 1000 units/ml of penicillin/streptomycin. Stable cell lines were selected by puromycin (MCF10A – 1 μM) or hygromycin (SUM149 – 150 μM, SUM159 – 300 μM, 231-LM2 – 700 μM, MCF10A – 100 μM) and the concentrations were empirically determined by killing curve assay. All resistant lines were pooled from surviving clones after 2-3 weeks selection. All trans-retinal (#R2500) and all trans-retinoic acid (#R2625) were purchased from Sigma.

2.2 Plasmid constructs

sh528 (TRCN0000017528) and sh529 (TRCN0000017529) were purchased from Open Biosystem. Tet-inducible shHOXA5 was generated by cloning targeting sequence from sh528 into pLKO-Tet-On plasmids purchased from Addgene. HOXA5 cDNA with myc-tag was cloned into pLHCX. Synonymous mutant (663bp to 683bp) was generated by site directed mutagenesis of HOXA5 for expressing sh529-resistant HOXA5 in MCF10A. For viral packaging, pLKO.1 vector (7 μg), pVSV-G (4 μg) and pCMV-dR8.91 (4 μg) or pLHCX vector (7 μg) and pCL-Ampho (7 μg) were transfected in 239T (10cm dish) for 36h and 72 hours prior to viral supernatant collection. Viral supernatant was filtered through 0.2 μm filter and 1ml of supernatant was used to infect cells in a 6-well plate. CDH1 and CD24 gene promoters were cloned into pGL2 luciferase plasmid.

2.3 Mammosphere formation assay

Tumor sphere assays with the breast cancer cell lines were performed as previously described (Dontu et al 2003) with modifications. For low density assay, single cell suspension from 0.25 x 10^4 of MCF10A-KRAS or 0.5 x 10^4 of SUM149 were seeded in 24-well ultra-low adhesion plates (Corning, #3473) in 1 ml of mammary epithelial growth media (MEGM, Lonza, #CC-3151), containing 1% methylcellulose (R&D Systems, #HSC001), to avoid aggregation and supplemented with 1X B27 (Invitrogen, #17504-44), 20 ng/mL EGF (Sigma, #E9644) and 20 ng/mL bFGF (Sigma, #F0291), and 4 μg/mL heparin (Sigma, #H1027). For secondary and tertiary sphere formation, spheres were dissociated into single cell at each passage by trypsinization (20 mins in 0.25% trypsin-EDTA at 37^{o}C and manually rock the tube few times) after growing for

14 days and then reseeded at $1x10^4$ cells/well. Quantification of the mammospheres was performed with ImageJ software.

2.4 3D growth assay

3D culture of the breast cancer cell lines was performed as previously described (Kenny et al 2007) with modifications. For MCF10A, 2 x 10^3 cells were inoculated into 100 µl growth factor reduced Matrigel (#354230, BD), plated in 8-well chamber slide (#155411, Nunc) and maintained in 400 µl assay medium. For MCF10A-KRAS, 2 x 10^3 cells were inoculated into 100 µl of 50% Matrigel (#354234, Corning) and 50% Collagen I (1 mg/ml final concentration and gelatinized according to manufacturer's protocol, #354236, Corning) mixture, plated in 8-well chamber slide and maintained in 400 µl of MCF10A complete media. For 231-LM2, 2.5 x 10^3 were inoculated into 100 µl growth factor reduced Matrigel (#354230, BD) plated in 8-well chamber slide and maintained in 400 µl 231-LM2 complete medium. Cells were allowed to grow for 12-16 days prior to imaging and enumerating colonies. Medium was replenished every 3-5 days.

2.5 Cell proliferation assay

Cells proliferation assay was performed as previously described (Korangath et al 2015). Cells were plated in 96 well plates at 1500-3000 cells per well in 100 µl media. Number of cells was determined by MTT colorimetric assay. 100 µl of 0.5 mg/ml MTT solution was added to medium depleted well for 3 hours. 200 µl of DMSO was added to dissolve the formazan crystals. This solution was read at 570 nm absorbance in a microplate reader (BMG Labtech).

2.6 BrdU incorporation assay

BrdU incorporation assay was performed according to manufacturer's instructions with modification. Five thousands cells were seeded in 96 wells for 24 hours before the addition of BrdU. BrdU was added for 3 or 18 hours prior measurement.

2.7 Wound healing assay

Cells were grown to confluence in 6-well plate at the time of assay. The monolayer was scored with a p200 pipette tip to create a scratch. Media was replenished and the scratch was allowed to close up for 20-40 hours. Representative images were taken and the remaining clear area was quantified by ImageJ software.

2.8 Cell Invasion assay

1.5-5 x 10^4 cells were seeded in 500 µl basal media in the Boyden chamber with 8 µm pores (#354480, BD) which was primed according to the manufacturer's protocol. Cells were incubated in the chamber for 16-20 hours and the Matrigel barrier was gently but thoroughly removed with a cotton swab. The membrane was fixed with 10% formalin (#HT501128, Sigma) for 30 minutes and the cells were stained with 0.05% crystal violet and de-stained in tap water for imaging. Three chambers were used for each test and cell numbers were determined from 4 independent fields in each chamber.

2.9 Quantitative RT–PCR

Total RNA was extracted using TRIzol (Invitrogen) and treated with DNase I. Reverse transcription was performed with 1 ug of total RNA using oligo-dT and MMLV reverse transcriptase (Promega), according to the manufacturer's instructions. Quantitative RT–PCR was performed with SYBR Green dye using ABI 7500 Fast Real-Time instrument

(Applied Biosystem). Relative expression level of each gene was calculated by the relative CT value compared to that of the internal control ACTB gene. Gel based RT-PCR was performed with optimal cycles and the PCR products were analyzed by 2% ethidium bromide agarose-gel. Primers (Sense, Anti-sense):

HOXA5 (5'-CCCCAGATCTACCCCTGGATG, 5'-ACGAGAACAGGGCTTCTTCA),
CDH1 (5'-CTGCTCTTGCTGTTTCTTCGG, 5'- TGCAGCTGGCTCAAGTCAAAG),
CLDN1 (5'-CGATGCTTTCTGTGGCTAAAC, 5'-
CTGAGGGCATCACTGAACAGA),

CD24 (5'-GGACATGGGCAGAGCAATGG, 5'-CCTTGGTGGTGGCATTAGTTG),
DSG3 (5'-CAAAGAAGTTCAGCCACCCTC, 5'-CTCCTTGACTTCCTGACAAAG),
ITGA10 (5'-CTAGGTGCCTGACTCCTTTC, 5'-GGAACAGGCAGGAAAGTGAG),
FOXA1 (5'-CTCTTTCCCCTCCAGACATTC, 5'-CAGACCTGTAAACTCGTAGGG),
FOXQ1 (5'-GCGGACTTTGCACTTTGAATCC, 5'-
CAAGAATTGCTTTGGTCCTCTG),

CDH2 (5'-CCCACACCCTGGAGACATTG, 5'-GAGCCACTGCCTTCATAGTC),
CDH3 (5'-CAGCCAACCCAGATGAAATCG, 5'-GCTGCCCTCATAGTCGAACAC),
CLDN7 (5'-CAGGCTATGGGAGTGTCTAG, 5'-ACCCAAGAGGACTATACATGG),
Occludin (5'-ATCAACTGGGCTGAACACTCC, 5'-
GTATGAAGACATCGTCTGGGG),

ACTB (5'- ATCAAGATCATTGCTCCTCCTG, 5'- CCTGCTTGCTGATCCACATC).

2.10 Western blot analysis

Cells were lysed with RIPA buffer containing protease and phosphatase inhibitor and mixed by rotation at 4^{o}C for 30 minutes. The lysates were centrifuged at 13,000 rpm for 20 minutes. The supernatant was collected and quantified by BCA assay. 20-50 μg of protein were used to run on SDS-PAGE gel after being heated with sample buffer. The proteins were then transferred to nitrocellulose membranes. After being blocked in 5% milk for 1 hour, the membranes were blotted with primary antibody overnight and with secondary horseradish peroxidase-conjugated antibody for 1 hour. The membranes were developed using chemiluminescence (Amersham) with or without 10% of enhancer (#1856190). Western blotting was using these antibodies: HOXA5 (#sc-365784, Santa Cruz); E-cadherin (#13-1700, Invitrogen); Claudin-1 (#71-7800), Claudin-7 (#34-9100), Occludin (#71-1500), ZO-1 (#33-9100) and CK18 (#180158Z) were from Zymed; p21 (Santa Cruz); β-actin (Sigma-A2228), cleaved caspase-3 (Cell Signaling).

2.11 Immunohistochemistry

IHC was performed as previously described (Korangath et al 2015) with modifications. Fresh xenograft tumors were fixed in 10% formalin and sectioned on positively charged slides. Vectastain ABC kit was used to perform IHC. After hydration with serial dilutions of ethanol, antigen retrieval was performed using 10 mM citrate buffer with 0.05% Tween-20. The sections were then treated with 3% hydrogen peroxide for 10 minutes and incubated with normal serum to block non-specific binding. The sections were then incubated overnight at room temperature with primary antibodies: HOXA5 antibody (1:50, # sc-365784, Santa Cruz), E-cadherin antibody (1:500, #13-1700, Invitrogen), CK-18 (1:500, #180158Z, Zymed). Secondary antibody was added the next

day after washing, followed by incubation with ABC reagent and developed with the DAB reagent, and counterstained with hematoxylin as specified by manufacturer.

2.12 Immunofluorescent staining and confocal microscopy

For 2D images, cells were fixed with methanol (E-cadherin) or 4% paraformaldehyde (Phalloidin and Occludin) and stained with corresponding antibodies using the manufacturer's protocol. Cells were treated with pre-extraction buffer prior to fixation for Occludin staining. Phalloidin-488 and secondary antibodies conjugated with Alexa-fluor 488 (Invitrogen) were used. Confocal IF staining of 3D organoids was performed as described (Debnath et al 2003). IF staining was using these antibodies: E-cadherin (#13-1700, Invitrogen), Phalloidin-488 or Phallodin-568 (both from Invitrogen), cleaved caspase-3 (Cell Signaling).

2.13 Fluorescence-Activated Cell Sorting and Quantification

One x 10^6 cells were stained on ice for 20 min with 5 µl of CD24-FITC (BD Pharmigen; clone ML5), 5 µl CD44-PE (BD Pharmigen; clone 515), 3 µl CD326 (EpCAM)-APC (Miltenyi Biotec, clone: HEA-125) and 5 µl of 7AAD (BD Pharmingen) to exclude dead cells, in 0.5% BSA/ 2mM EDTA/ PBS. Samples were run on the BD FACSCalibur system (Becton Dickinson), and data analyzed using FlowJo software. For apoptosis assay, 1 x 10^6 cells were stained with 2.5 µl of Annexin V-Alexa 647 (Life Technologies) and 5 µl of 7ADD. Cells were sorted on a BD FACSAria SORP and analyzed on a BDLSRII with BD FACSDiva Software (BD Biosciences).

2.14 Xenograft and Limiting Dilution Assay

For the limiting dilution assay, breast cancer cells were digested with 0.05% Trypsin-EDTA. Viable single cells were determined with Trypan blue exclusion and serially diluted to achieve 0.25 x 10^6-0.25 x 10^3 cells per injection. Cells were mixed with Matrigel and injected into mammary fat pads of immunocompromised NOD/SCID interleukin-2 receptor gamma chain null [Il2rg(-/-)], NOG mice. The fat pads were humanized by injecting a mixture of irradiated and non-irradiated (1:1) immortalized human fibroblast cells (25,000 cells/100µl Matrigel/fat pad) (Ginestier et al 2007, Kuperwasser et al 2004). Skin fibroblasts were provided by J. Campisi (Krtolica et al 2001).

2.15 Promoter Luciferase assay

Confluent **(**90%) 239T cells in 12-well plate were co-transfected with 1µg amount of corresponding vectors, 50 ng of promoter luciferase constructs and 20 ng of β-galactosidase plasmid using Lipofectamine 2000 (Invitrogen). The luciferase activity (#E1501, Promega) was measured 24 hours after transfection of cells and normalized to β-galactosidase activity (#117758241001, Roche).

2.16 ChIP assay

Methods described in (Jin et al 2012) were followed with modifications. Ninety percent confluent cells on a 10 cm dish was cross-linked with 1% formaldehyde and then quenched with 0.125 M glycine. Scraped and pelleted cells were lysed and sonicated using optimized conditions using a Bioruptor sonicator. Sonicated chromatin was enriched by using either 1 µg of α-myc or 3 µg of α-HOXA5 and same amount of IgG

isotype control and column-purified. ChIP primers (Sense, Anti-sense) used for CDH1 were (5'-AAAATTAGGCCGCTCGAGGC, 5'-GTATTTTTAGTACAGACGGGGT), and for CD24 were (5'-AGAGTGATGACGAATCTCCTC, 5'-CCCGTCATTCGTTTAAAGAGC). Distance sites primers used were: upstream CD24 (5'- ATTGTAAGCATTGCTCCTCCC, 5'- GAGAAAGTGTGGTATAGAATGA), downstream CD24 (5'- TGCGAGGCAACCTTTCACGT, 5'-GTAGCGGCGACCCGAACAAA), upstream CDH1 (5'-GGGACAGAATTCTCTGGAGAA, 5'-CAAAGCCAAAGACCAAAACCC), downstream CDH1 (5'- GAAGAGGAGGTTGAGGGCAC, 5'-CAAAGCCAAAGACCAAAACCC).

2.17 Gene expression array and analysis

RNA was extracted from MCF10A cells with RNeasy mini kit (#74104, Qiagen) according to manufacturer's protocols. RNA was obtained from biological duplicate of each: MCF10A scramble, MCF10A-sh528 and MCF10A-sh529. RNA was hybridized onto Illumina's HumanHT-12v4 BeadChip (Illumina, San Diego) using manufacturer recommended protocols. These chips provide comprehensive coverage of >47,000 transcripts, including all current NCBI curated genes. After pre-processing using Illumina GenomeStudio software, quantile normalization was applied to log transformed intensities for future analyses. Statistical analyses were performed and figures of array data generated on the R statistical software using standard R tools, packages from the Bioconductor Bioinformatics Software Project, and customized programs when necessary.

2.18 Statistical Analysis

Quantitative data were expressed as mean ± standard deviation (SD) of experimental triplicates or otherwise stated in the legends, and all statistical tests were analyzed by two-tailed student's t test or otherwise stated, and considered statistically significant at $p<0.05$ [GraphPad Prism (v5.0, GraphPad Software, San Diego, CA)]. Tumor growth was presented as mean ± SEM and the statistical analysis performed using two tailed Student t test. Gene set enrichment analysis (GSEA) was performed with GSEA software on expression array data using default parameters on canonical pathway collections (Subramanian et al 2005). Statistical analysis and boxplot comparing HOXA5 expression across grade were generated from breast cancer data sets (Desmedt et al 2007, van 't Veer et al 2002) with R statistical software using standard R tools. Significance testing was performed using Welch's corrected t-tests with Bonferroni correction for multiple testing. Relapse free survival curve using HOXA5 as a marker was performed with a web-based tool-Kaplan Meier plotter using median as the cut off (Gyorffy et al 2010). Tumor initiation frequency for the limiting dilution assay in vivo was determined by Poisson distribution evaluated by L-Calc software (with 95% confidence).

3 Results

3.1 HOXA5 expression is frequently lost in breast cancer.

Previous studies have provided evidence for loss of HOXA5 expression in breast cancer cell lines and in primary tumors (Coradini et al 2014, Henderson et al 2006, Raman et al 2000). Analysis of TCGA dataset showed that HOXA5 level is significantly lower in breast cancer compared to normal breast tissue, regardless of breast cancer subtype (Figure 3.1-1). MCF10A, a non-transformed but immortalized mammary epithelial cell line, was used as a model to decipher the biological functions of HOXA5 in vitro. MCF10A, among other immortalized

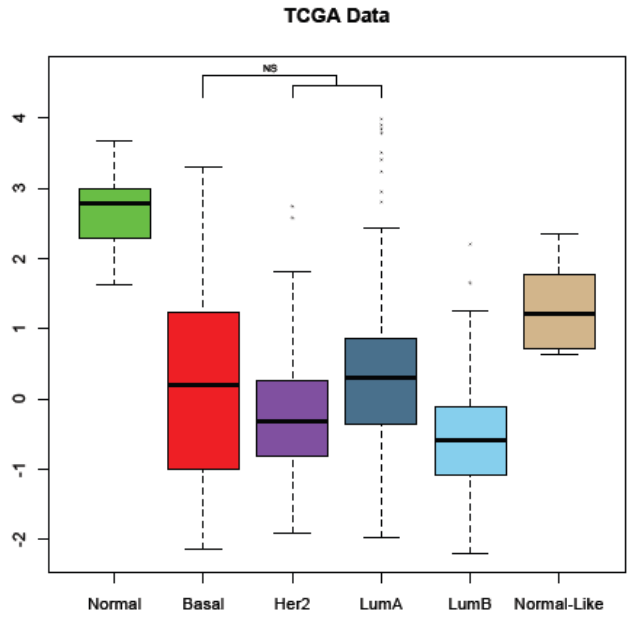


Figure 3.1-1. *HOXA5 expression level of different breast cancer sub-group in TCGA dataset. Pairwise comparisons of each group are statistically significant except the indicated pairs (NS= not significant).*

breast cell lines, expressed a level of HOXA5 similar to primary human mammary epithelial cells (HMEC) while the majority of breast cancer cell lines express low or undetectable levels of endogenous HOXA5 (Figure 3.1-2).

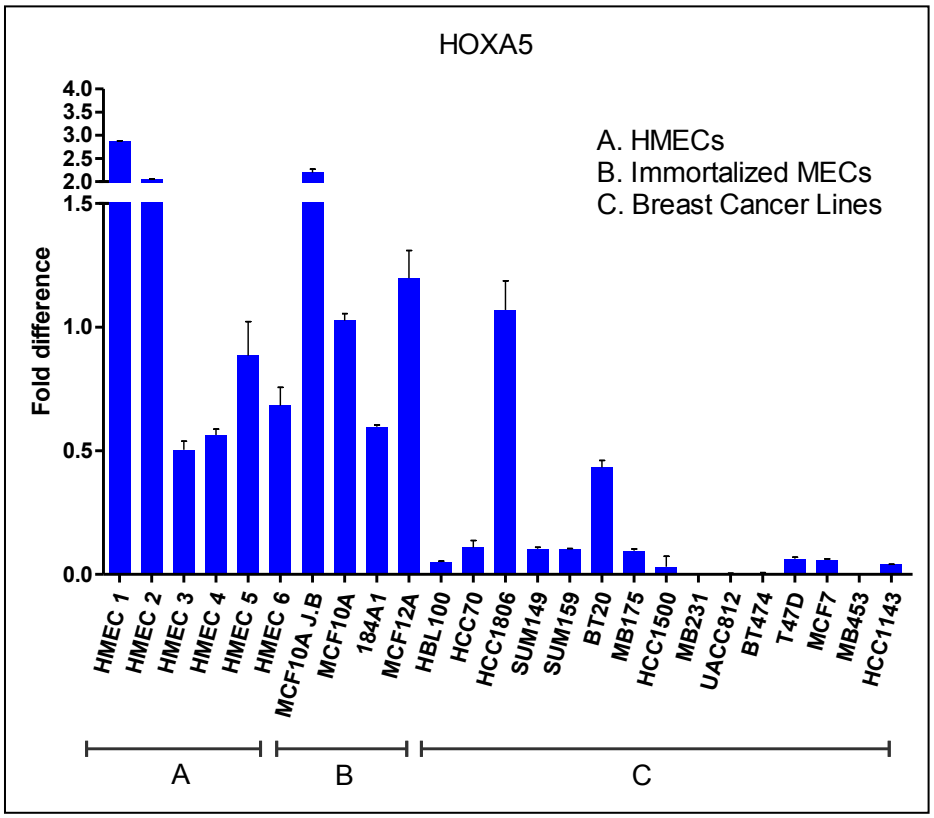


Figure 3.1-2. *HOXA5 mRNA expression in HMECs and breast cell lines.*

We confirmed the loss of HOXA5 protein in breast cancer by performing immunohistochemistry (IHC) on primary breast carcinomas. HOXA5 nuclear staining was detected in histologically normal mammary glands within the same section (Figure 3.1-3A). We observed a progressive loss of HOXA5 from histological normal mammary gland to invasive ductal carcinoma lesions on the same section (Figure 3.1-3B). Further, analysis of the expression profiles comparing micro-dissected cancer and normal lesions from same section revealed that reduction in HOXA5 expression is

indeed epithelial specific (Figure 3.1-4). These findings using clinical samples indicate that the loss of HOXA5 may contribute to the progression of breast cancer.

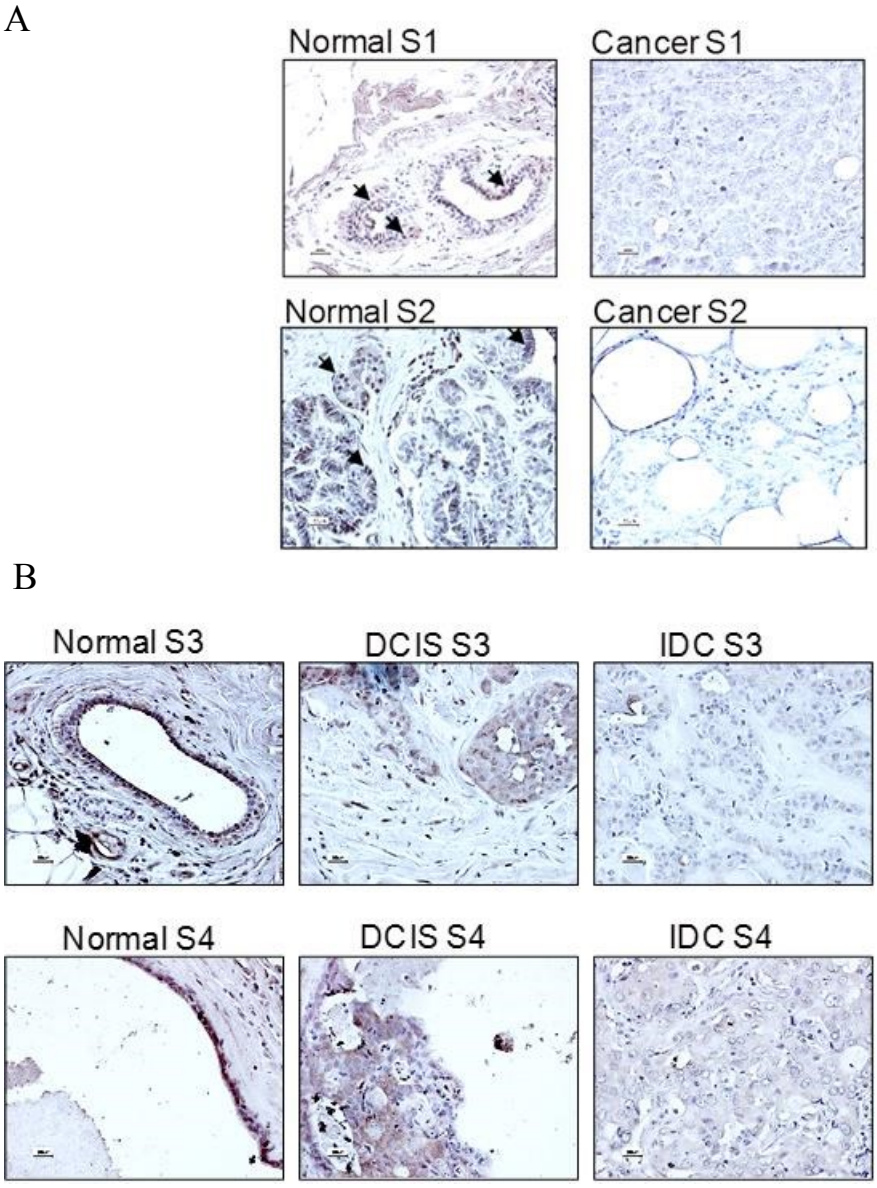


Figure 3.1-3. *Immuno-histochemistry of breast cancer for HOXA5. A. Representative IHC images for HOXA5 of histological normal mammary glands (Normal) and invasive ductal carcinoma (Cancer). B. Representative IHC images for HOXA5 of histological normal mammary glands (Normal), ductal in situ carcinoma (DCIS) and invasive ductal carcinoma (Cancer). Arrows indicate nuclear staining of HOXA5.*

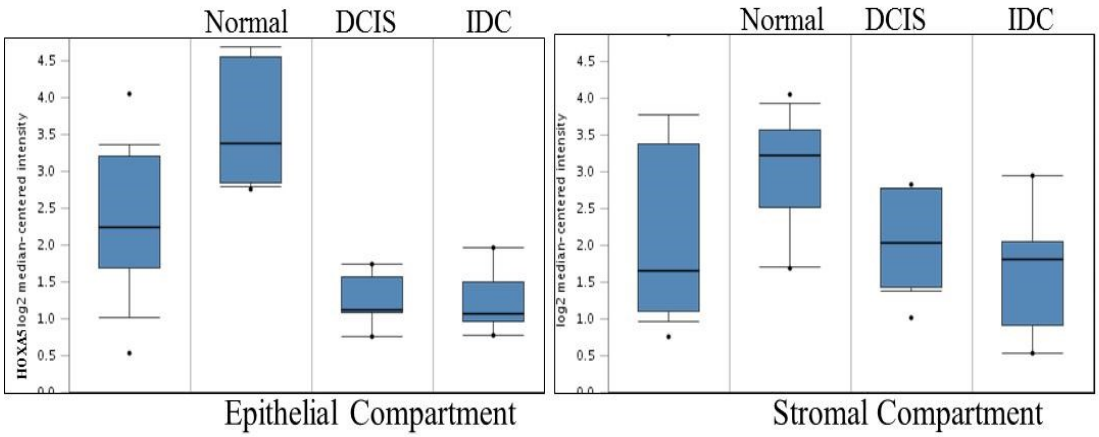


Figure 3.1-4. *HOXA5 expression in micro-dissected normal epithelium, normal stroma, tumor epithelium and tumor-associated stroma (left panel: epithelial compartment; right panel: stromal compartment.*

3.2 Depletion of HOXA5 perturbs epithelial characteristics.

We depleted HOXA5 in MCF10A with a set of 5 shRNA and evaluated the expression of HOXA5 to determine the level of reduction. Two (sh528 and sh529) out of five shRNA were successful in substantially reducing the levels of HOXA5 protein (Figure 3.2-1A) and mRNA (Figure 3-2-1B). Further, as shown by q-RT-PCR, the level of depletion of HOXA5 in these sh-clones was more robust and consistent over several culture passages (Figure 3.2-1C).

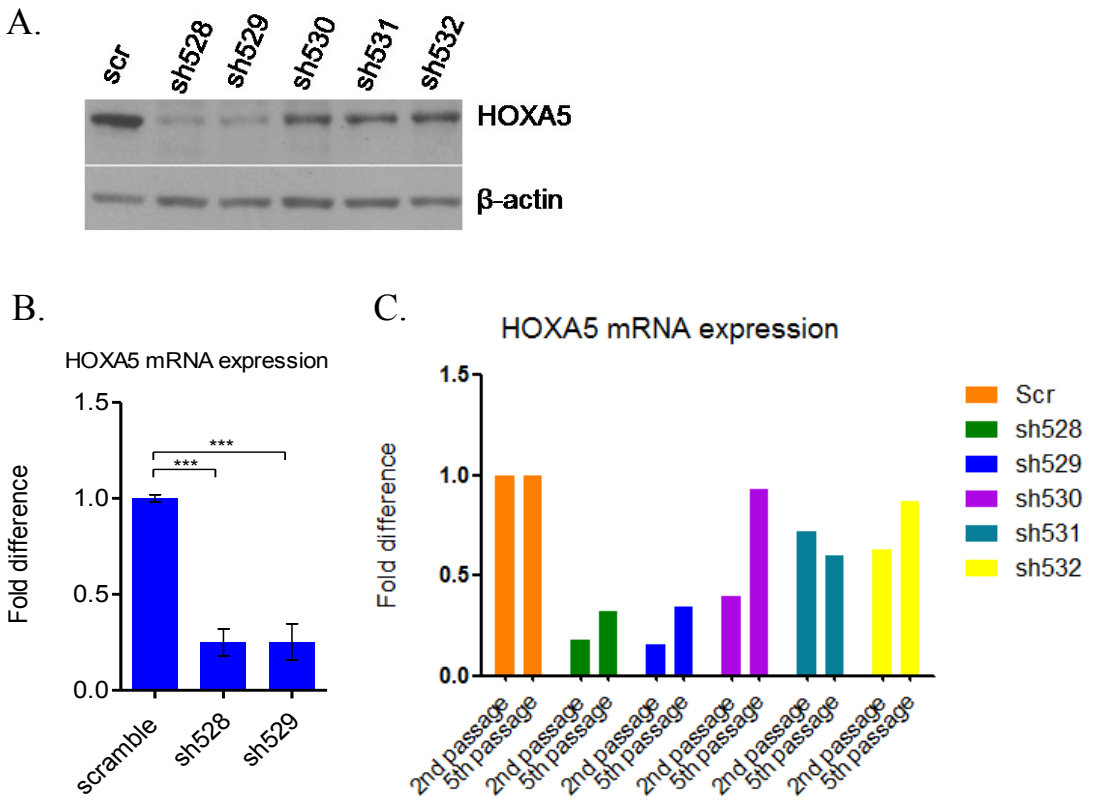


Figure 3.2-1. *Depletion of HOXA5 in MCF10A cells. A. Western blot analysis of endogenous HOXA5 in MCF10A cell lines stable express scramble shRNA, HOXA5 targeting sh528, sh529, sh530, sh531 and sh532. B. RT-PCR analysis for HOXA5 mRNA level of MCF10A cell lines stable express scramble, sh528 and sh529 shRNAs and C. HOXA5 targeting sh528, sh529, sh530, sh531 and sh532 over two time points of passage. In both assays, HOXA5 were substantially and robustly depleted with sh528 and sh529 compared to scramble control.*

To gain further insight into pathways through which HOXA5 functions as a tumor suppressor gene, microarray analysis was performed with MCF10A cultures, MCF10A-sh528 and MCF10A-sh529, depleted of HOXA5 using two shRNAs (Figure 3.2-1). Pairwise differential expression analysis performed using Limma across

biological replicates identified 43 genes, including HOXA5 itself, that are differentially regulated with an FDR < 0.01 (Figure 3.2-2). The pipeline resulted in a subset of differentially expressed genes that were consistently and robustly altered across MCF10A-sh528 and MCF10A-529 compared to MCF10A-scr control.

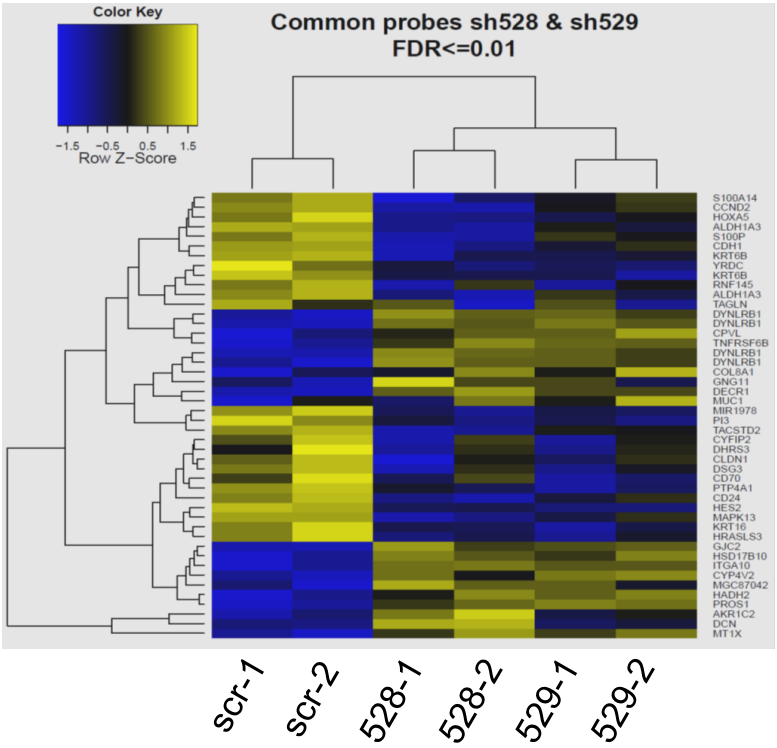


Figure 3.2-2. *Hierarchical clustering heatmap showing genes differentially expressed between MCF10A-scr cells (n=2, biological duplicates) and HOXA5 depleted -sh528 and -sh529 cells (n=2 each, biological duplicates) that have an FDR < 0.01.*

We performed quantitative RT-PCR to validate a subset of highly differentially expressed genes identified by microarray analysis including CDH1, a cell adhesion protein; CD24, a differentiation marker; ALDH1A3, involved in retinal biosynthesis; CLDN1, a tight junction protein; ITGA10 and DSG3 (Figure 3.2-3A). To confirm the direct involvement of HOXA5 in the expression of these genes, HOXA5 was transfected back into these cells. In both HOXA5-depleted cell cultures, ectopic expression of HOXA5 led to a substantial recovery of mRNA expression of each of the genes lost as a result of silencing HOXA5 (Figure 3.2-3A), suggesting that loss of gene expression was a direct consequence of HOXA5 loss. The data also showed that HOXA5 expression correlated with the expression of multiple genes that are necessary for the maintenance of epithelial cell identity in MCF10A.

We also examined the expression level of several enzymes that are associated with retinoic acid biosynthesis in HOXA5-depleted MCF10A cells. The expression of AKR1C2 and AKR1B1 (members of aldo-keto reductase), as also of DHRS3, were decreased in HOXA5-depleted MCF10A cells. In contrast, proteins that have been implicated in enhanced cell invasion including ANPEP and MMP2 were elevated. An neuroendocrine secretory protein, CHGB, was also increased in HOXA5-depleted MCF10A cells. Ectopic expression of HOXA5 in the knock down cells, MCF10A-sh528 and MCF10A-sh529 cells reversed the expression pattern of these genes (Figure 3.2-3B).

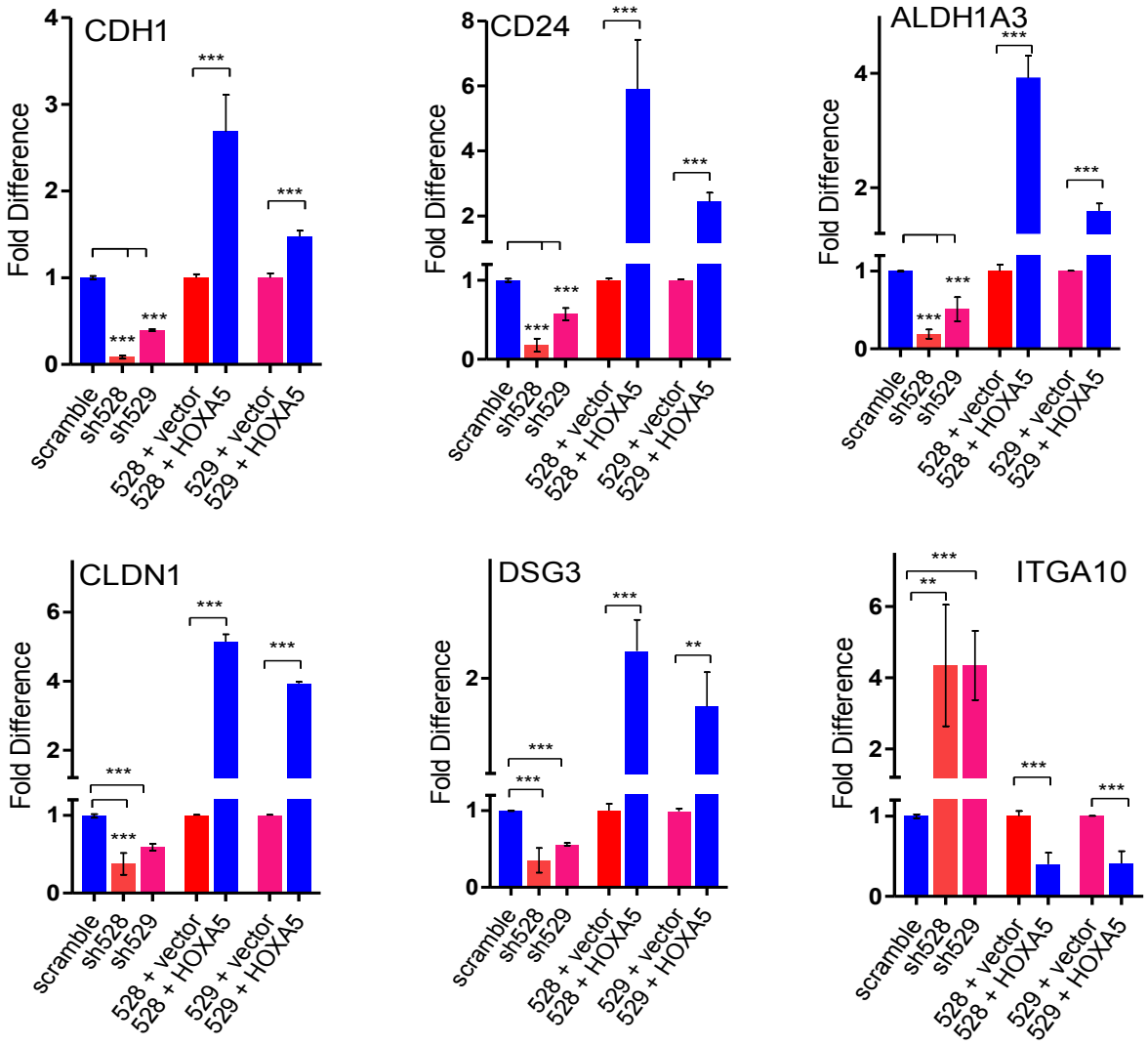


Figure 3.2-3. *Validation of the genes identified by micro-array analysis. A. Quantitative RT-PCR validation of a selection of five underexpressed (CDH1, CD24, ALDH1A3, CLDN1 and DSG3) and one overexpressed gene (ITGA10) identified in HOXA5-shHOXA5 cells by the array.*

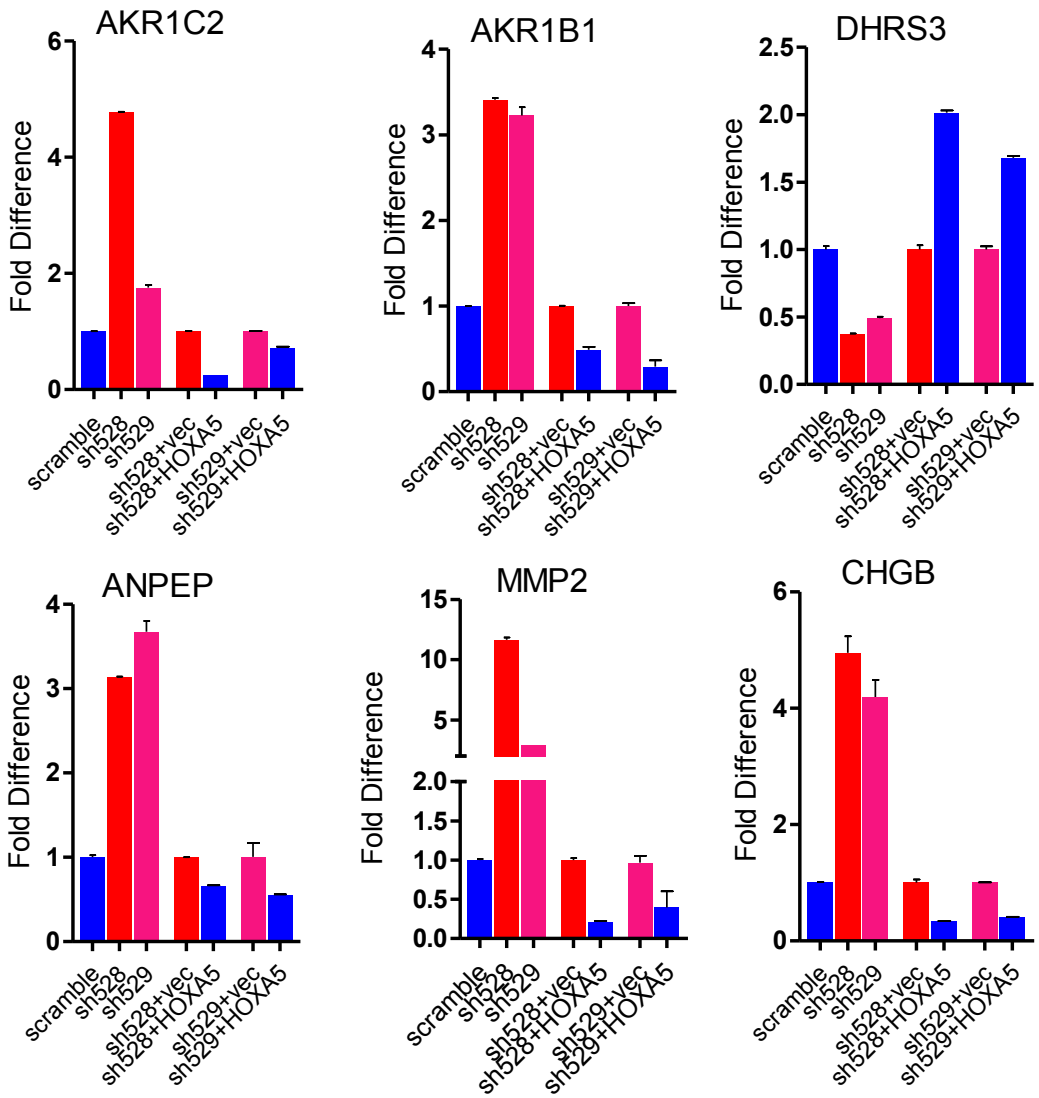
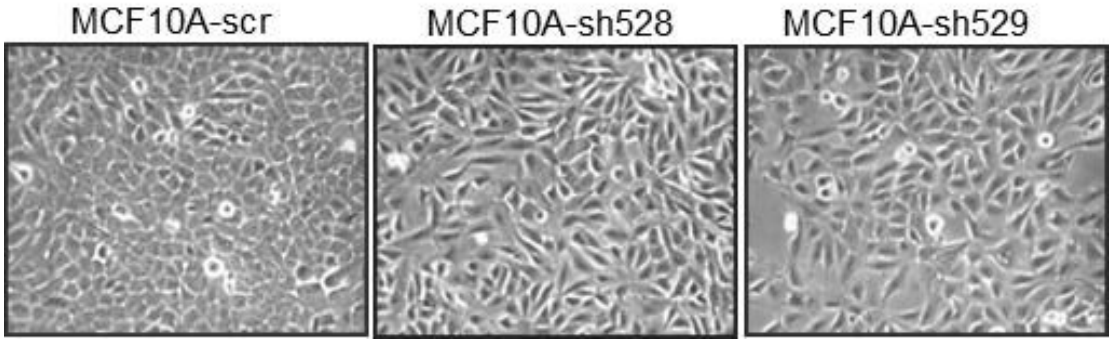


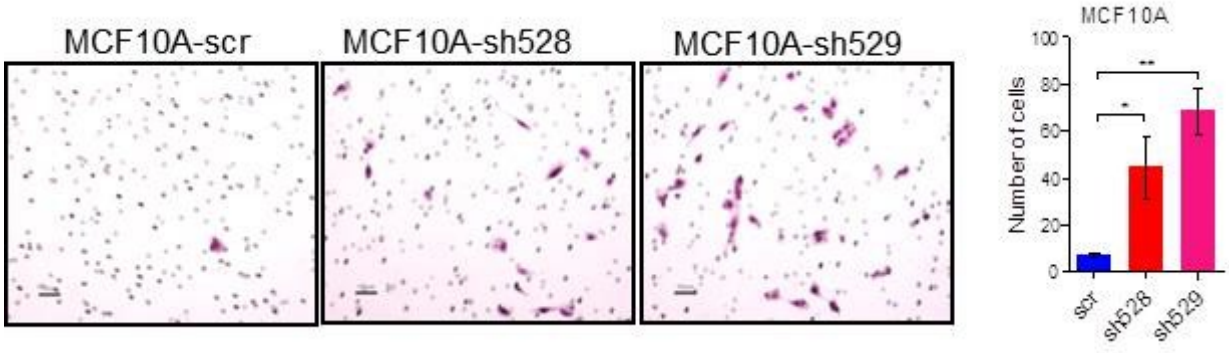
Figure 3.2-3 *B. Quantitative RT-PCR of genes discovered through microarray analysis that are involved in retinoic biosynthesis (AKR1C2, AKR1B1 and DHRS3) and cell invasion (ANPEP, MMP2 and CHGB).*

One easily observable change that occurred in the phenotype of both MCF10A-sh528 and MCF10A-sh529 cell lines was a loss of typical cuboidal epithelial morphology (Figure 3.2-4A), which was accompanied by a significantly increased capacity to invade through Matrigel barrier in Boyden chambers (Figure 3.2-4B) Conversely, ectopic expression of HOXA5 in aggressive SUM149, SUM159 and LM2 [a lung metastatic derivative of MDA-MB-231 (Padua et al 2008)] breast cancer cells resulted in strikingly impeded invasiveness (Figure 3.2-4C-E). Furthermore, depletion of HOXA5 in breast cancer model MCF10A-KRAS cells also increased its invasiveness (Figure 3.2-4F).

A.



B.



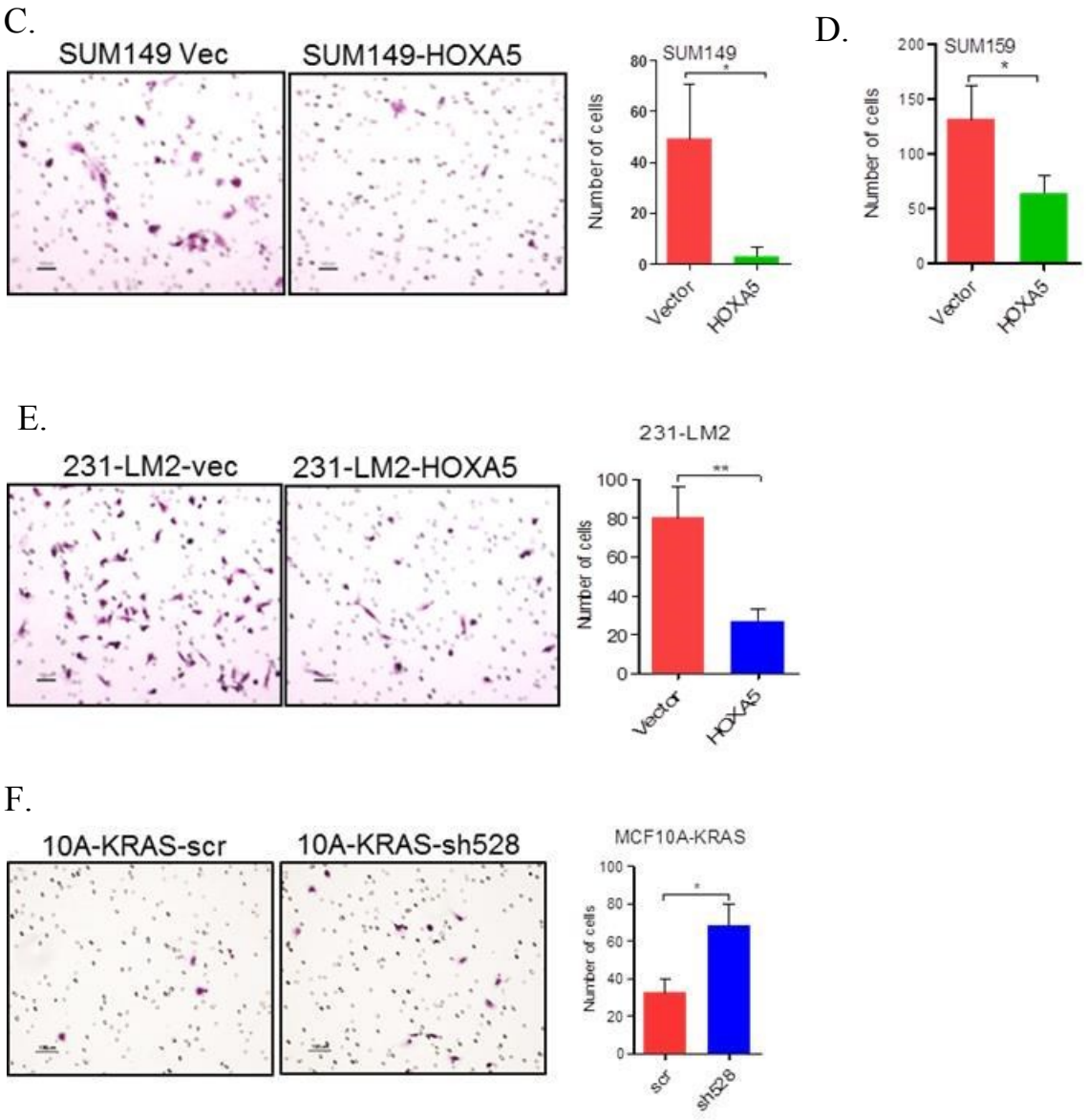
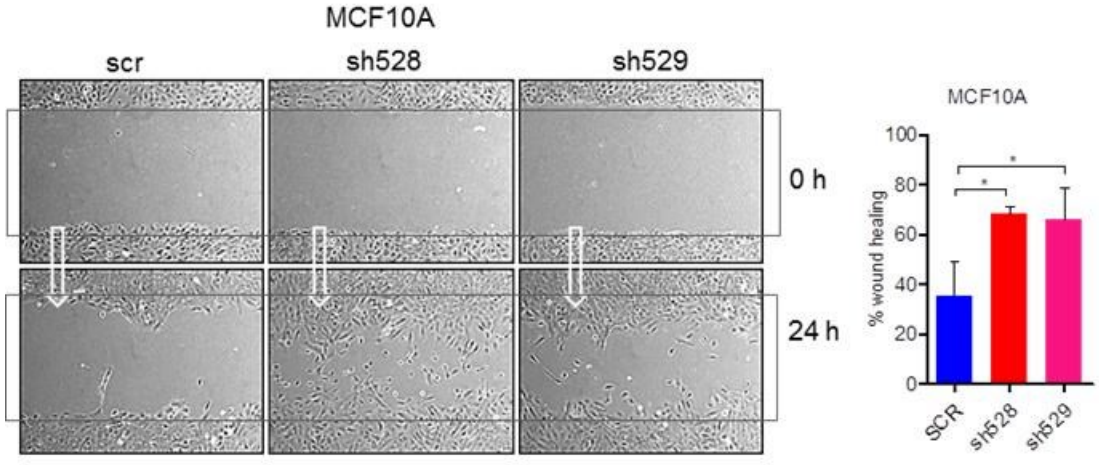


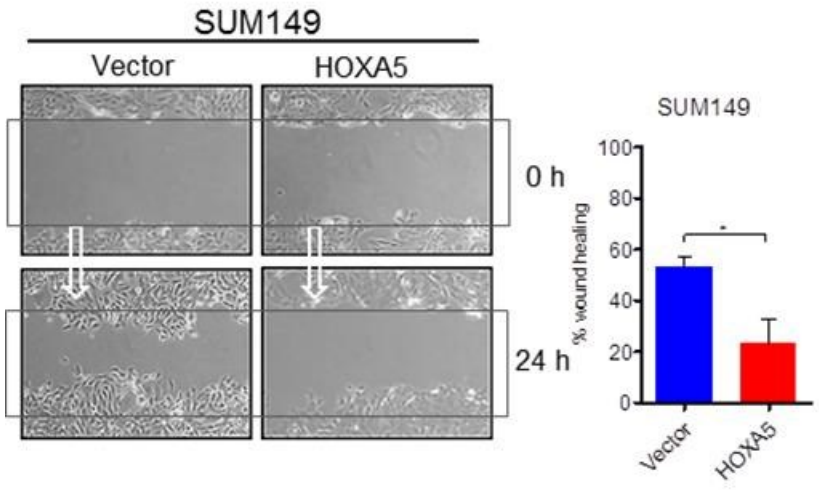
Figure 3.2-4. *Loss of HOXA5 increases cell invasiveness. A. Representative phase-contrast images of MCF10A-scr and HOXA5 depleted -sh528 and -sh529 cells. Representative images and quantitative analysis of invaded cells in Boyden chamber and quantitative analysis of number of invaded cells in: B. MCF10A-scr and HOXA5 depleted -sh528 and -sh529 cells. C. SUM149-vector and SUM149-HOXA5 cells. D.SUM159-vector and SUM159-HOXA5 (Quantitative analysis). E. MDA-MB-231-LM2-vec, and MDA-MB-231-LM2-HOXA5 cells. F. MCF10A-KRAS-scr and MCF10A-KRAS–sh528 cells. Quantitation is presented in bar graphs.*

Depletion of HOXA5 also altered the motility of MCF10A cells. In line with the observed traits, MCF10A-sh528 and –sh529 cells showed a two-fold higher efficiency in wound healing/migration capability compared to MCF10A-scr cells (Figure 3.2-5A), while the converse was evident in SUM149-HOXA5, SUM159-HOXA5 and LM2-HOXA5 cells (Figure 3.2-5, B, D, E). Depletion of HOXA5 in MCF10A-KRAS model also slightly increased the motility of the cells (Figure 3.2-5C).

A.



B.



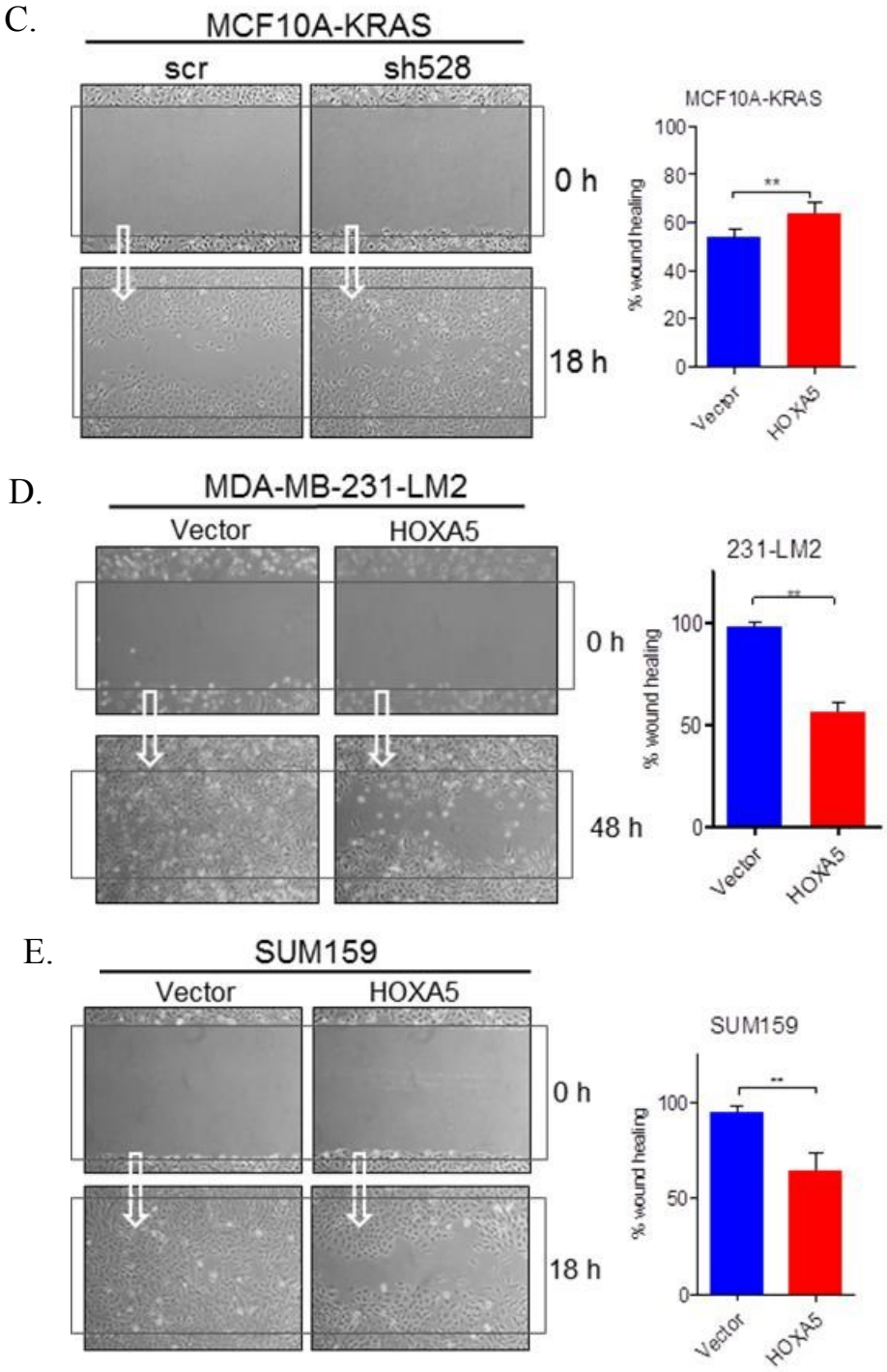


Figure 3.2-5. *HOXA5 impedes breast cancer cells motility. Representative images of wound healing assay and quantitative analysis of percent wound healing over different time course: A. MCF10A-scr and HOXA5 depleted -sh528 and -sh529 cells (24 hours). B. SUM149-vector and SUM149-HOXA5 cells (24 hours). C. SUM159-vector and*

SUM159-HOXA5 cells (18 hours) D. MDA-MB-231-LM2-vec and MDA-MB-231-LM2-HOXA5 cells (48 hours). E. MCF10A-KRAS-scr and MCF10A-KRAS–sh528 cells (18 hours). Quantitation is presented in bar graphs.

In four out of the five cells lines, the changes in cell invasion and cell motility were not dictated by altered cell proliferation, since no significant difference in cell proliferation was noted between controls and HOXA5-depleted or HOXA5-overexpressing cells (Figure 3.2-6A-E). However, ectopic expression of HOXA5 significantly reduced the growth of 231-LM2 cells. We further examined cell proliferation by performing BrdU incorporation assay. Similarly, no significant change in cell proliferation was observed between controls and HOXA5-depleted or HOXA5-overexpressing cells (Figure 3.2-7A-E).

All these pieces of evidence support the notion that HOXA5 expression influences the phenotype and behavior of the cell lines. In most cases, HOXA5 expression status does not alter cell proliferation and growth. Together with the microarray analysis, the data is highly suggestive of the fact that HOXA5 regulates epithelial traits, which in turns maintain the typical characteristics of epithelial cells.

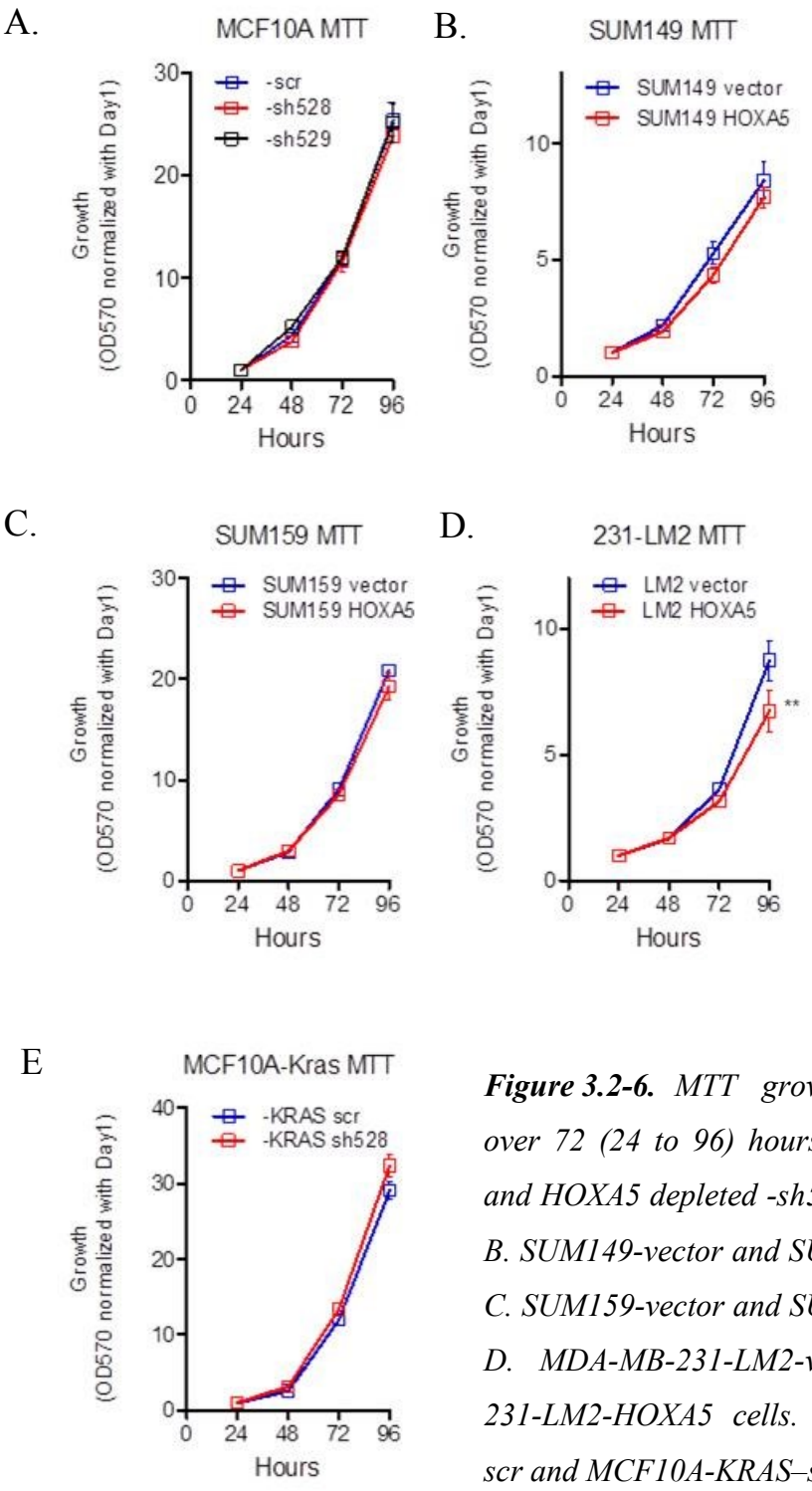


Figure 3.2-6. *MTT growth analysis (n=5) over 72 (24 to 96) hours of A. MCF10A-scr and HOXA5 depleted -sh528 and -sh529 cells. B. SUM149-vector and SUM149-HOXA5 cells. C. SUM159-vector and SUM159-HOXA5 cells. D. MDA-MB-231-LM2-vec and MDA-MB-231-LM2-HOXA5 cells. E. MCF10A-KRAS-scr and MCF10A-KRAS–sh528 cells.*

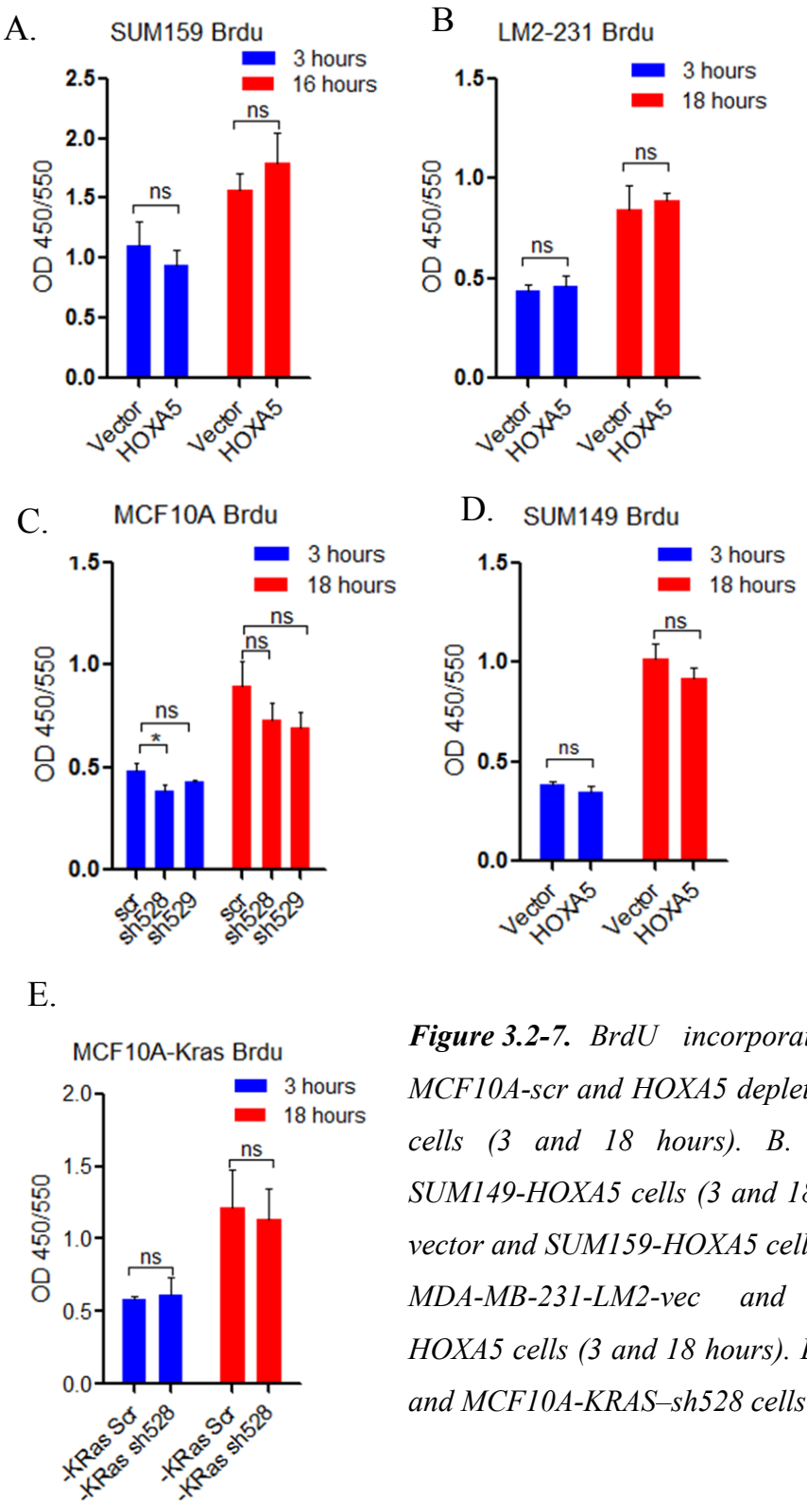


Figure 3.2-7. *BrdU incorporation analysis of A. MCF10A-scr and HOXA5 depleted -sh528 and -sh529 cells (3 and 18 hours). B. SUM149-vector and SUM149-HOXA5 cells (3 and 18 hours). C. SUM159-vector and SUM159-HOXA5 cells (3 and 16 hours). D. MDA-MB-231-LM2-vec and MDA-MB-231-LM2-HOXA5 cells (3 and 18 hours). E. MCF10A-KRAS-scr and MCF10A-KRAS–sh528 cells (3 and 18 hours).*

3.3 Depletion of HOXA5 perturbs cell-cell adhesion.

Some of the key molecular changes identified through our microarray analysis were those associated with epithelial traits of cell-cell adhesion and interaction. Gene set enrichment analysis (GSEA) of our microarray data revealed that MCF10A-scr cells were enriched for genes involved in cell-cell-adhesion and tight junctions (Figure 3.3-1).

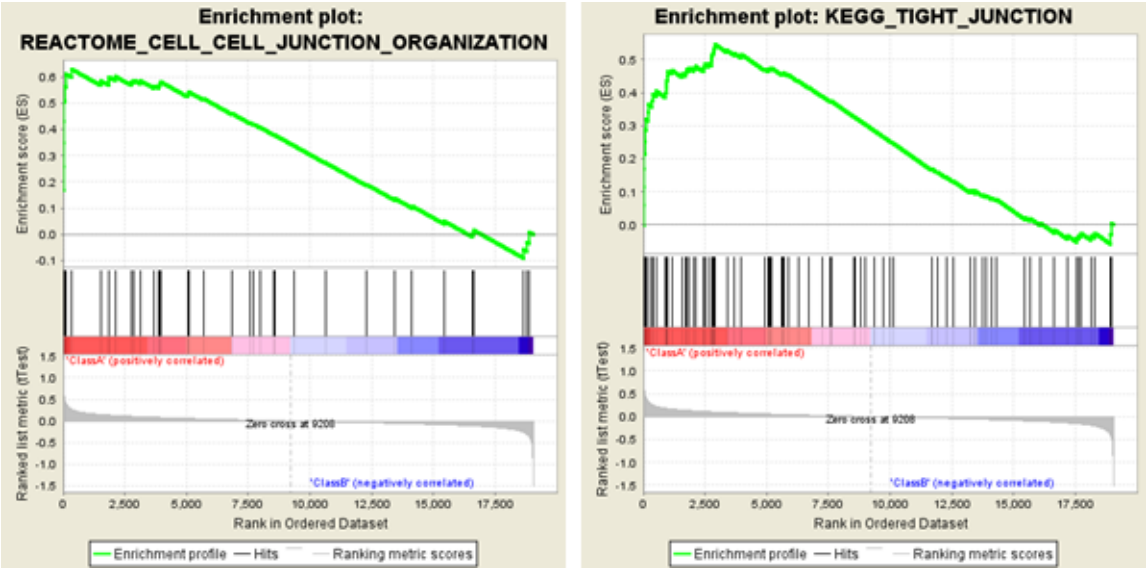


Figure 3.3-1. *Gene set enrichment analysis of MCF10A-scr and HOXA5 depleted cells indicate that cell-cell junction genes are enriched in MCF10A-scr cells. Gene set enrichment analysis (GSEA) was performed with GSEA software on expression array data using default parameters on canonical pathway collections. (Class A: MCF10A-scr, Class B: MCF10A-shHOXA5).*

The analysis corroborated our findings on the compromised epithelial traits observed in HOXA5-depleted MCF10A cells. Further, we examined a subset of genes that are associated with epithelial differentiation of mammary cells. Analyzing these

molecular changes associated with compromised epithelial integrity by RT-PCR, we observed loss of expression of E-cadherin (CDH1) was accompanied by an elevation of N-cadherin (CDH2) in HOXA5 depleted cells (Figure 3.3-2). Also evident was the decrease of expression of P-cadherin (CDH3) (Figure 3.3-2), which is known to disrupt the normal suppressor function of CDH1 by decreasing the interaction between CDH1 and intracellular catenins (Ribeiro and Paredes 2014). Occludin and claudin-7 (CLDN7), integral components of tight junctions in mammary epithelial cells, were also down-regulated upon HOXA5 depletion (Figure 3.3-2). These molecular changes suggested that the loss of function in adhesion junction and tight junctions was associated with the phenotypic switch. Two other regulators of EMT, acting in opposing ways, were also investigated. There was loss of expression of the transcriptional factor FOXA1, a member of the forkhead class of DNA-binding proteins, which has been shown to potently neutralize several EMT-related, CDH1-repressed mechanisms (Wang et al 2013), while there was overexpression of another member of the same family, FOXQ1 (Figure 3.3-2), which

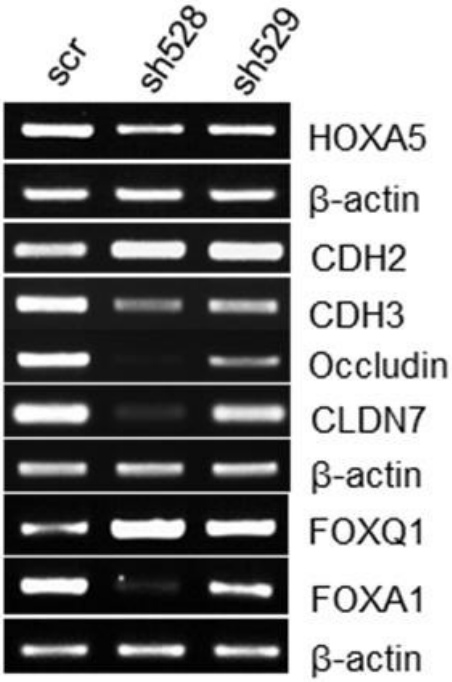


Figure 3.3-2. *RT-PCR analysis of differentiation markers in MCF10A-scr, HOXA5 depleted -sh528 and -sh529 cells including EMT markers (CDH2 and CDH3), tight junction molecules (Occludin and CLDN7) and luminal fate markers (FOXA1 and FOXQ1).*

represses CDH1 by binding to the E-box in its promoter region (Qiao et al 2011, Zhang et al 2011).

As detected by Western blotting, the loss of epithelial traits in HOXA5-depleted MCF10A cells was accompanied by significant reduction in protein levels of CDH1, occludin, CLDN7 and CLDN1 (Figure 3.3-3).

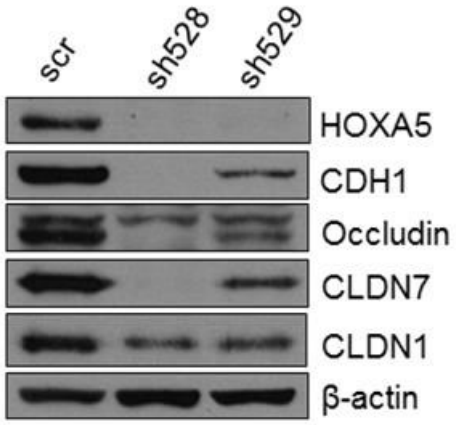


Figure 3.3-3. *Western blot analysis of differentiation markers in MCF10A-scr, HOXA5 depleted -sh528 and -sh529 cells including HOXA5 and HOXA5-regulated CDH1, Occludin, CLDN7 and CLDN1. β-actin serves as the loading control.*

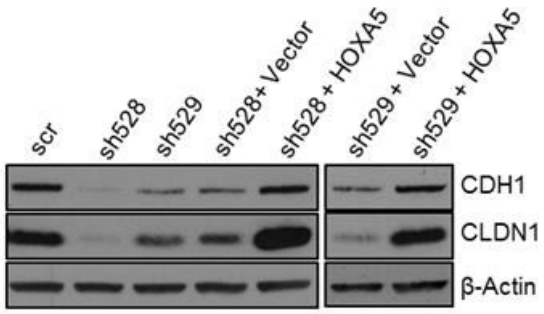


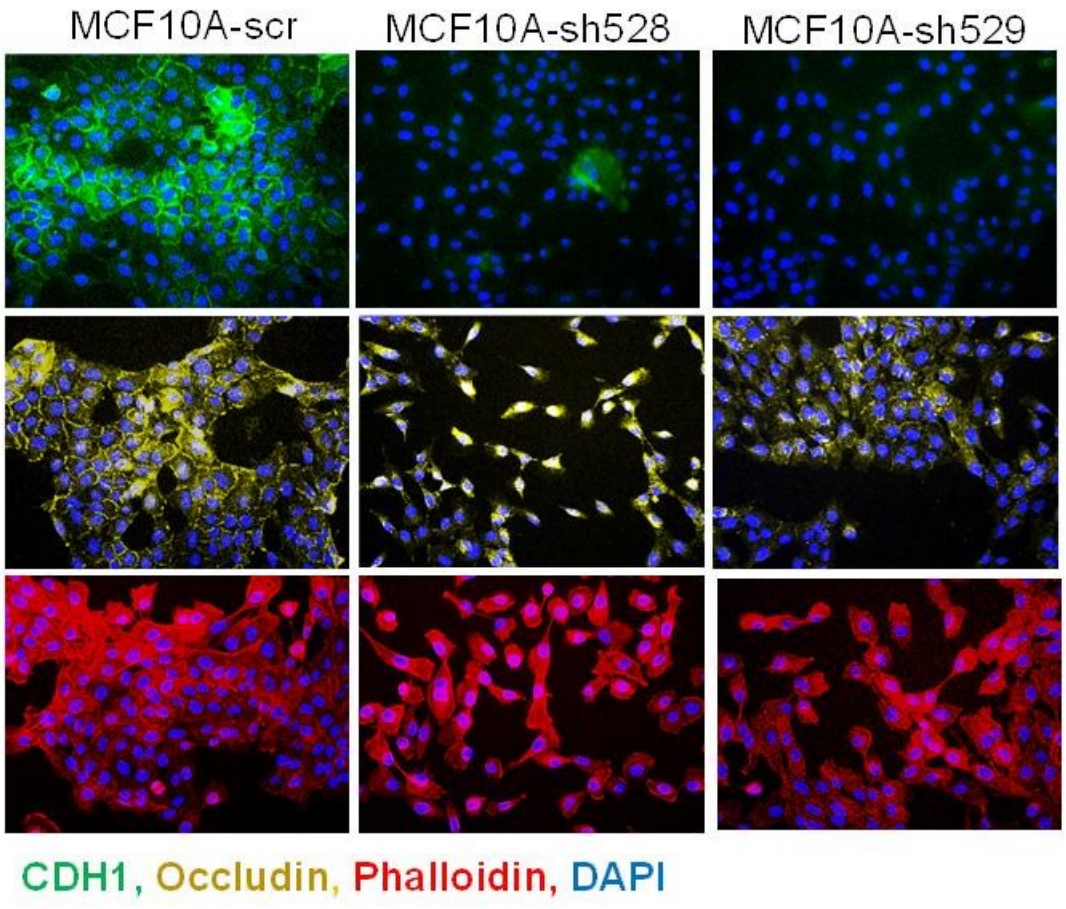
Figure 3.3-4. *HOXA5 regulates CDH1 and CLDN1 in MCF10A cells. Western blot analysis shows decrease in CDH1 and CLDN1 protein upon HOXA5 depletion in MCF10A cells; and reversal by ectopic expression of HOXA5 into the knock-down clones. β-actin serves as the loading control.*

Furthermore, analysis of stable depletion of HOXA5 in MCF10A-sh528 and MCF10A-sh529 cells markedly reduced CDH1 and CLDN1 expression, while re-expression of HOXA5 in these cells restored CDH1 and CLDN1 expression (Figure 3.3-4) and hence suggested a direct regulatory role of HOXA5 on CDH1 and CLDN1. HOXA5 thus

appears to promote cell-cell adhesion and expression of epithelium-specific markers in breast cells.

Immunofluorescence (IF) analysis provided additional evidence for membrane loss of CDH1 and Occludin, conferring an elongated morphology to the HOXA5-depleted MCF10A cells (Figure 3.3-5A). The loss of membrane bound adhesion molecules upon the loss of HOXA5 in MCF10A cells was evident (Figure 3.3-5B). The analyses of the typical epithelial markers revealed the regulation of multiple components of adhesion and tight junctions by HOXA5, and thus strengthen the differentiation role in HOXA5.

A.



B.

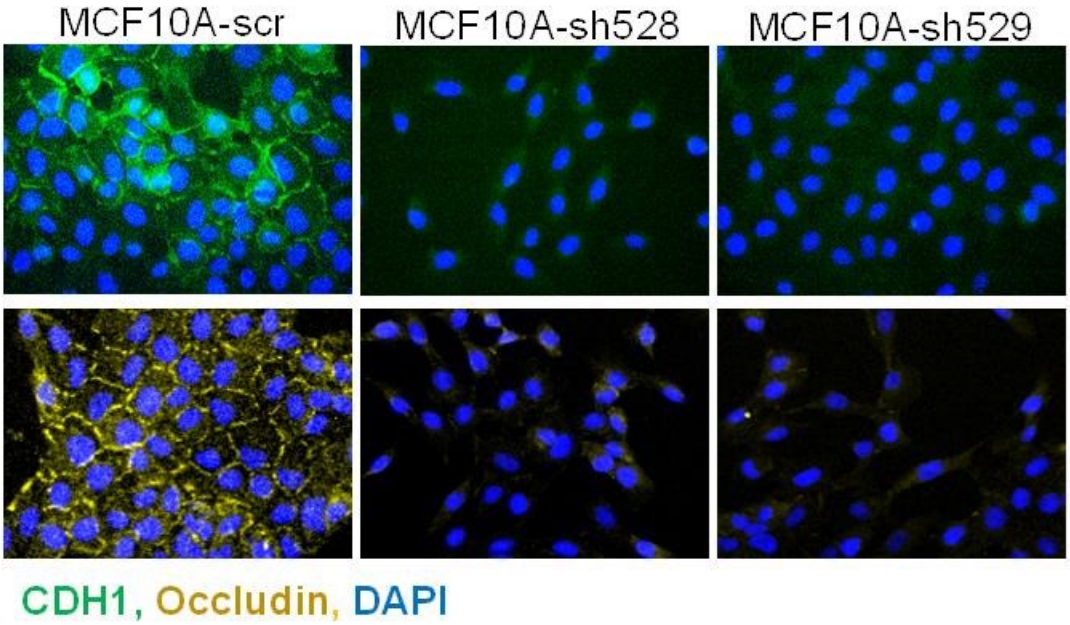


Figure 3.3-5. *HOXA5 regulates membrane-bound junctional proteins. A. Representative immunofluorescence images of CDH1 (Green), Occludin (Yellow), actin stress fiber, Phalloidin (Red) and nuclear marker, DAPI (Blue). B. Representative immunofluorescence images at 40X magnification of CDH1 (Green) and Occludin (Yellow).*

3.4 HOXA5 expression inhibits outgrowth of organoids in three dimensional (3D) cultures.

Transformed breast epithelial cells form protrusive and microinvasive structures in Matrigel, strongly associated with the epithelial-mesenchymal transitional state of the cells (Kenny et al 2007, Qiao et al 2011). Since depletion of HOXA5 resulted in the loss of epithelial properties in MCF10A, MCF10A-sh528 and MCF10A–sh529 cells were cultured in three-dimensional (3D) in Matrigel to assess their growth patterns.

Depletion of HOXA5 in MCF10A led to the formation of organoids of a more complex structure in Matrigel culture (Figure 3.4-1).

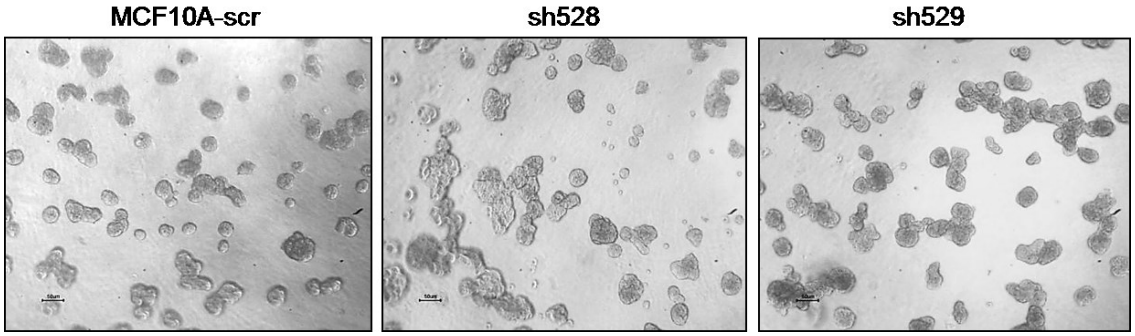


Figure 3.4-1. *Depletion of HOXA5 results in abnormal organoids formation. Representative phase-contrast images of MCF10A organoids formed in matrigel 3D culture.*

231-LM2 breast cancer cells express low or undetectable level of CD24 and Claudin-1. CD24 and Claudin-1 expression were restored when HOXA5 was ectopically expressed (Figure 3.4-3A). Since the proliferation of HOXA5 expressing LM2 cells was significantly reduced, we performed flow cytometry analysis using Annexin V and 7ADD staining to determine the apoptosis of the cells. 231-LM2-HOXA5 cells exhibited an increased proportion of early apoptotic cells (Q4) and late apoptotic cells (Q3) (7.3% vs 10.6% and 4.5% vs 9% respectively, Figure 3.5-3C) although the increase was subtle. Nevertheless, cleaved caspase-3 was not detected in either 231-LM2-vector or –HOXA5 cells (Figure 3.4-3B), suggesting that cells did not undergo rigorous apoptosis in culture. In 3D assays in Matrigel, MDA-MB-231 lung metastatic derivative, 231-LM2-vec cells formed multiple branching and stellate structures in two weeks, while outgrowth of HOXA5 overexpressing 231-LM2-HOXA5 cells was significantly inhibited (Figure 3.4-3A). Phalloidin staining of the organoids

corroborated the restricted growth phenotype in 231-LM2-HOXA5 cells (Figure 3.4-3B, quantified in 3.4-3C). We further performed cleaved caspase-3 staining in 3D culture to confirm the state of 231-LM2-HOXA5 cells in 3D culture. Cleaved caspase-3 staining was more prominent in the spheroid structure, which is the dominant structure in HOXA5 expressing 231-LM2 cells in matrigel. These structures also exhibited increased apoptotic blebs which is characterized by de-nucleated and GFP+ cytoplasm (Figure 3.4-3D).

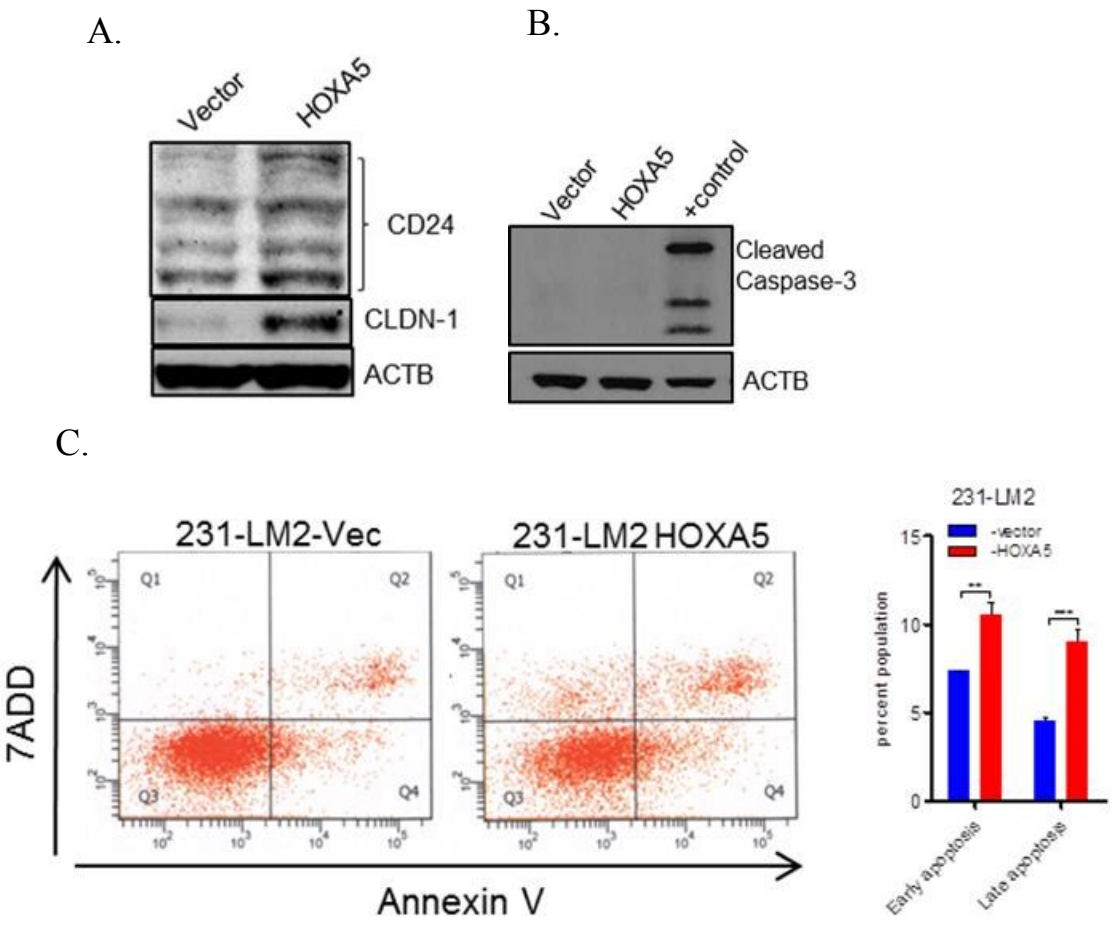


Figure 3.4-2. *HOXA5 increases CD24 and CLDN1 but not cleaved caspase 3 in 231-LM2 cells. Western blot analysis for A. CD24 and CLDN1 B. Cleaved caspase-3 in 231-LM2-vector or 231-LM2-HOXA5cells. . β-actin serves as the loading control. C. Flow*

cytometry analysis of Annexin V (x-axis) against 7ADD (y-axis) in 231-LM2-vector and 231-LM2—HOXA5 cells. Quantitation is shown in the bar graphs.

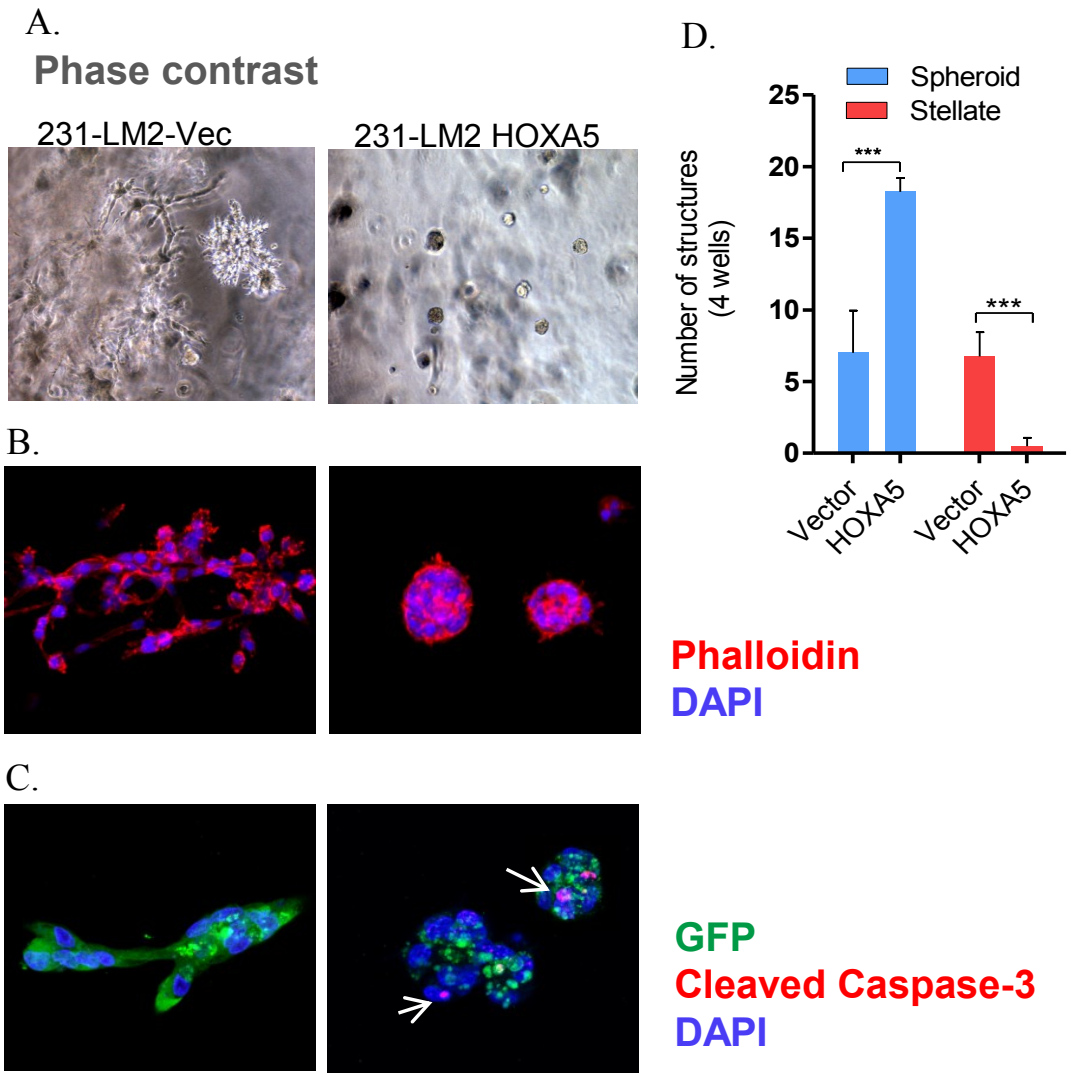


Figure 3.4-3. *HOXA5 restricts organoids outgrowth and promotes apoptosis of 231-LM2 cells in matrigel. Representative A. phase contrast and B. confocal images showing vector or ectopic HOXA5-expressing 231-LM2 organoids visualized with Phalloidin (Red) and DAPI (Blue). C. confocal images showing vector or ectopic HOXA5-expressing 231-LM2 organoids visualized with cleaved caspase-3 (Red), GFP (Green) and DAPI (Blue). D. Quantitative analysis of spheroids or branched structures formed in Matrigel.*

On the other hand, depletion of HOXA5 in MCF10A-Kras cells resulted in protruding outgrowths into a matrigel/collagen gel (Figure 3.4-4A-B) and this was accompanied by the loss of CDH1 expression (Figure 3.4-4C) and a significant increase in the number of branches (Figure 3.4-4D). Collectively, the data showed that expression of HOXA5 impedes protrusive and microinvasive outgrowth of organoids in 3D cultures, and the increase in CD24 and Claudin 1 protein expression are supportive of HOXA5-mediated differentiation of breast cancer cells.

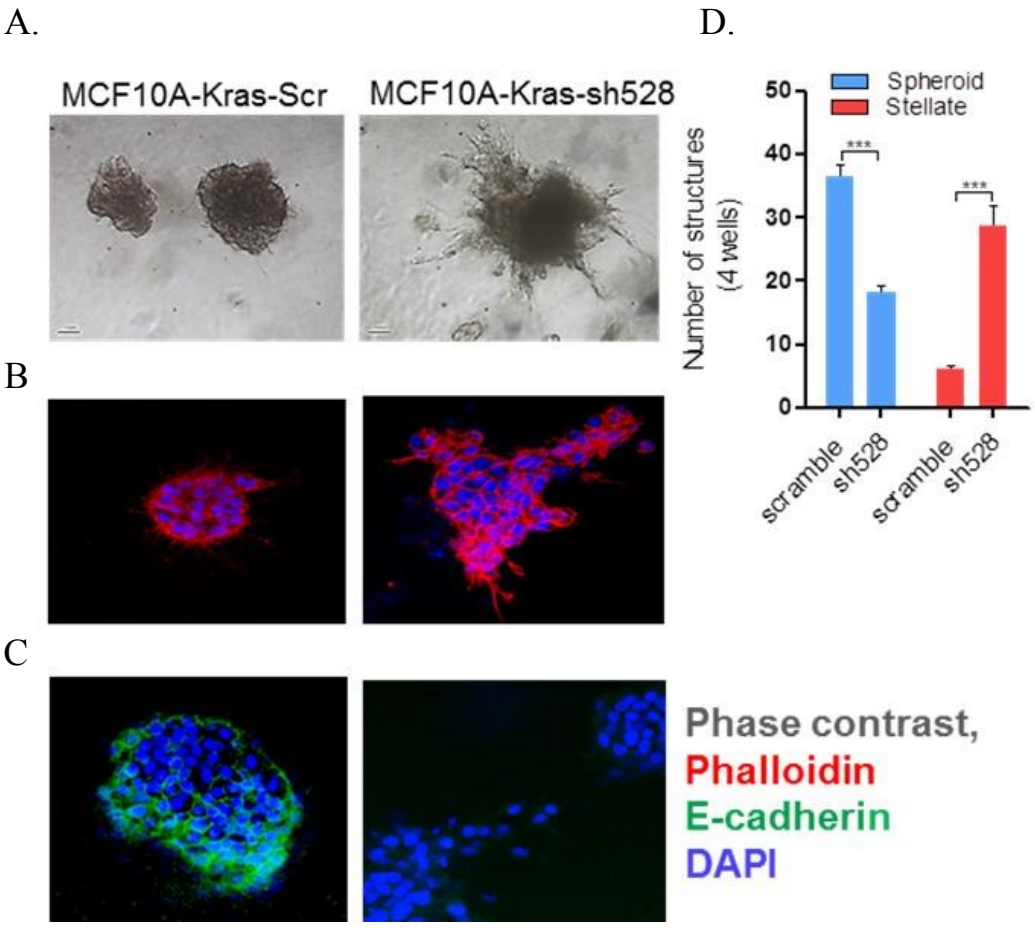


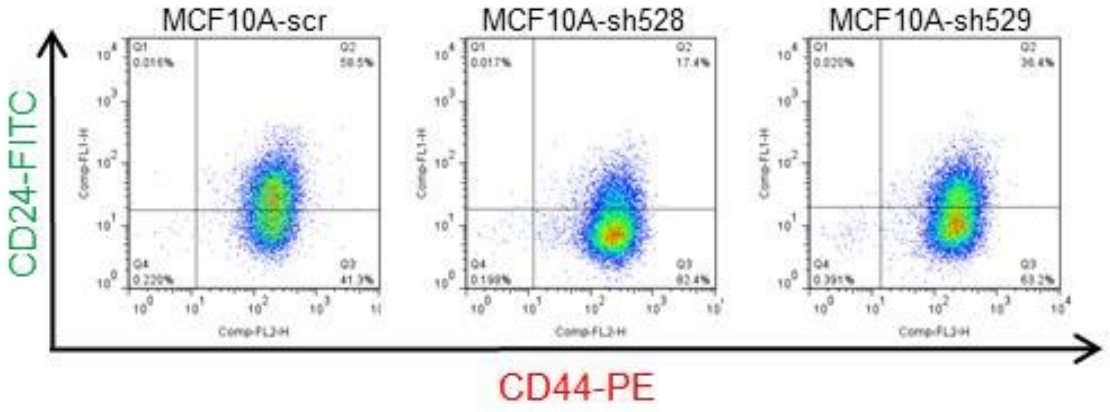
Figure 3.4-4. *HOXA5 prevents organoids outgrowth in MCF10A-Kras cells. Representative A. phase contrast, and confocal images of MCF10A-Kras-scr or HOXA5-depleted –sh528 organoids grown in Matrigel: Collagen I matrix and stained*

with B. Phalloidin (Red), C. CDH1 (Green) and DAPI (Blue). D. Quantitative analysis of spheroids or branched structures of MCF10A-Kras cells formed in Matrigel/collagen.

3.5 HOXA5 loss leads to an increase in stemness of epithelial cells.

Human mammary epithelial cells have often been shown to acquire stem-like characteristics when they undergo EMT in vitro (Mani et al 2008, Radisky and LaBarge 2008). Alterations typical of EMT, observed in our studies, raised the possibility of similar changes brought about by depletion of HOXA5 in MCF10A cells. Flow-cytometry analysis of HOXA5-depleted MCF10A-sh528 and -sh-529 cells provided evidence for a reduction of the $CD24^{+}/CD44^{+}$ (from 65.8% to 20.2% and 42.9% respectively) population compared to MCF10A-scr cells (Figure 3.5-1A-B). There was an increase in the number of mammospheres formed by MCF10A-sh528 and –sh529 cells (Figure 3.5-1C), suggesting that loss of HOXA5 skewed the MCF-10A cell population to a more progenitor-like state.

A.



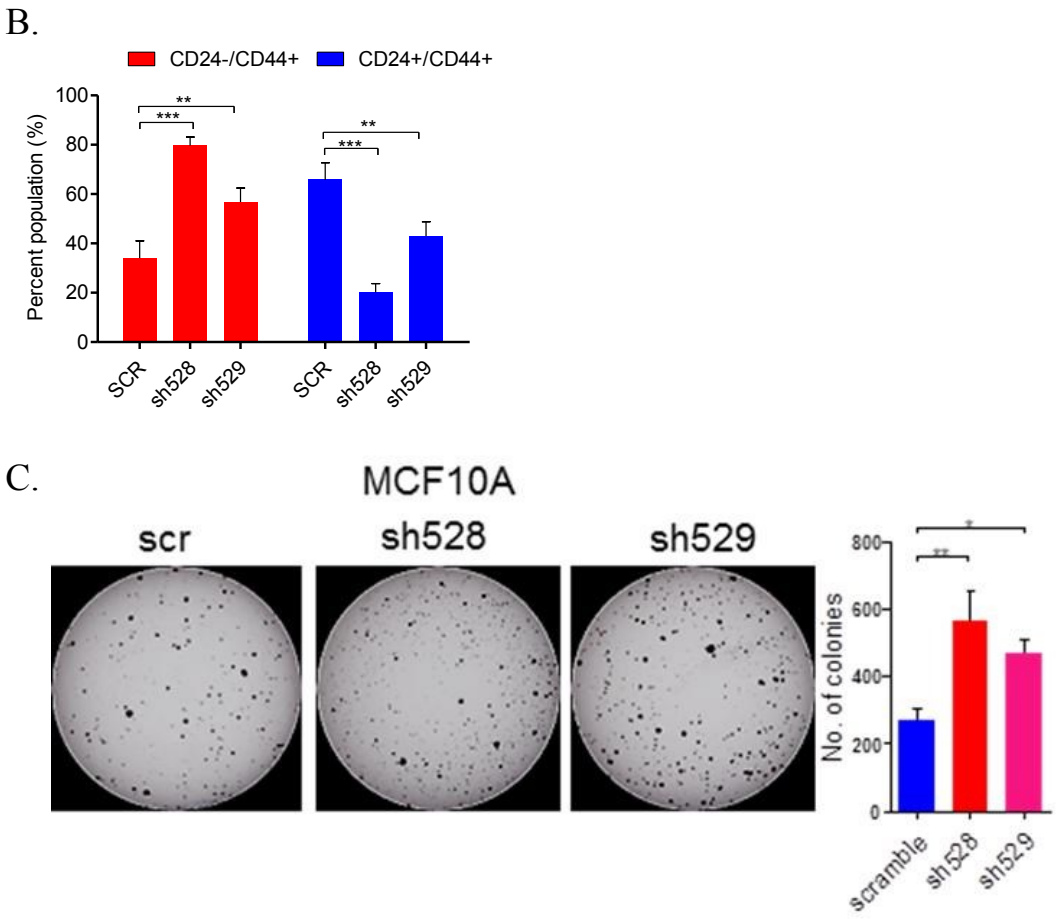


Figure 3.5-1. *Depletion of HOXA5 enhances cell stemness in MCF10A. A. Flow cytometry analysis of CD24 (x-axis) against CD44 (y-axis) in MCF10A-scr, HOXA5 depleted -sh528 and -sh529 cells. B. Quantitation is shown in the bar graphs. C. Representative images and quantitative analysis of mammosphere formation by MCF10-scr, HOXA5 depleted -sh528 and -sh529 cells.*

Our findings that loss of HOXA5 in MCF10A cells lead to increased migration, invasion, epithelial-mesenchymal transition, both in phenotypic and molecular characteristics, and acquisition of a more progenitor cell properties supported a role for HOXA5 as a differentiation factor in normal breast cells. This led us to investigate its

functions in MCF10A cells stably transformed by the K-ras oncogene. Depleting HOXA5 with shRNA in MCF10A-Kras significantly decreased the number of $CD24^{+}/CD44^{lo}$ cells (from 22.7% to 3%) (Figure 3.5-2A-B) and increased the number of mammospheres formed in culture (Figure 3.5-2C).

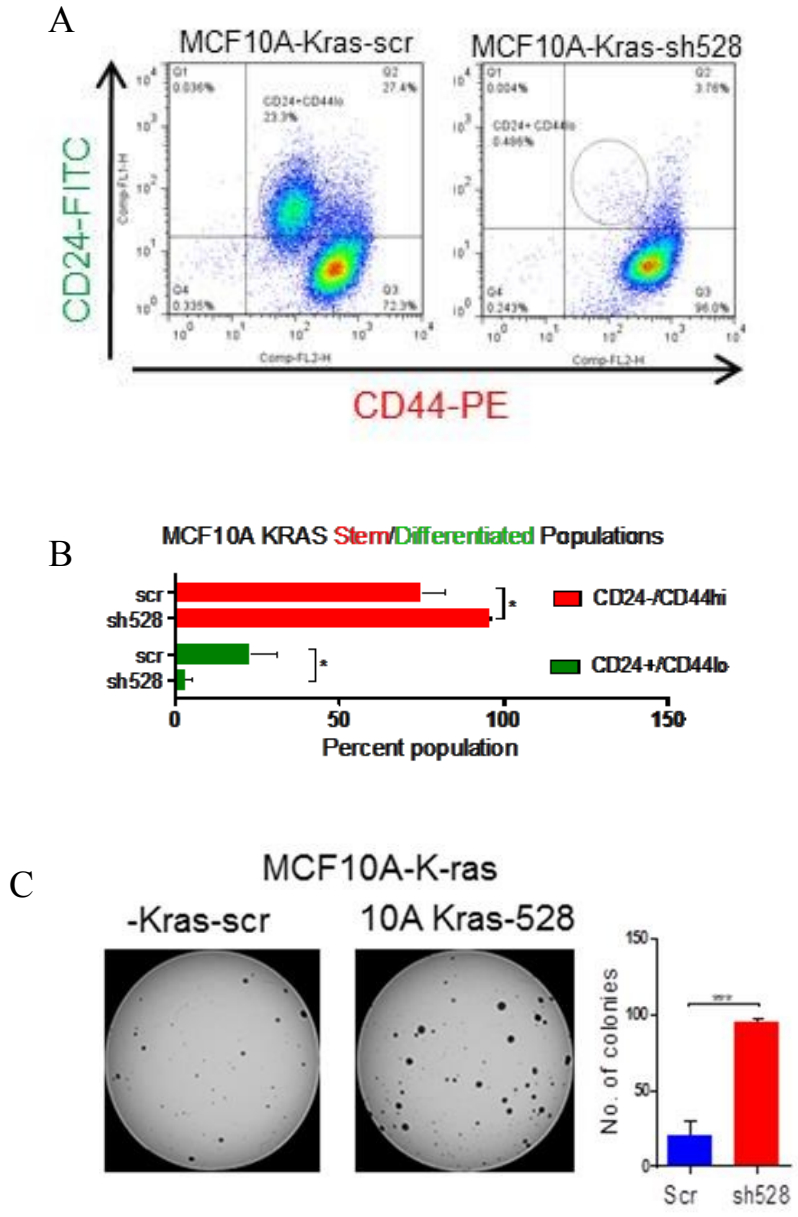
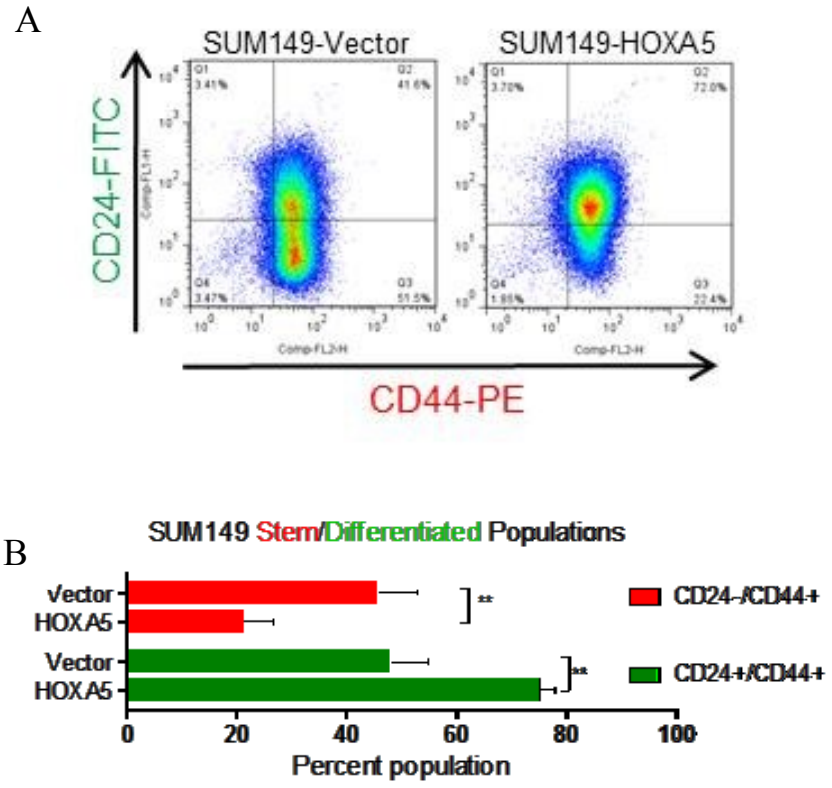


Figure 3.5-2. *HOXA5 inhibits cell stemness in MCF10A-Kras. A. Flow cytometry analysis of CD24 (x-axis) against CD44 (y-axis) in MCF10A-scr, HOXA5 depleted -*

sh528 and -sh529 cells. B. Quantitation is shown in the bar graphs. C. Representative images and quantitative analysis of mammosphere formation by MCF10-scr, HOXA5 depleted -sh528 and -sh529 cells.

Conversely, in HOXA5 over-expressing SUM149 cells we observed an increased $CD24^{+}/CD44^{+}$ (from 47.9% to 75.1%) population (Figure 3.5-3A-B) and reduced mammosphere formation (Figure 3.5-3C) compared to SUM149-vector control. Over-expressing HOXA5 in SUM159 did not alter mammosphere formation; however $CD24^{+}/CD44^{+}$ population was elevated by 30% when analyzed by flow-cytometry compared to vector control (Figure 3.5-3D).



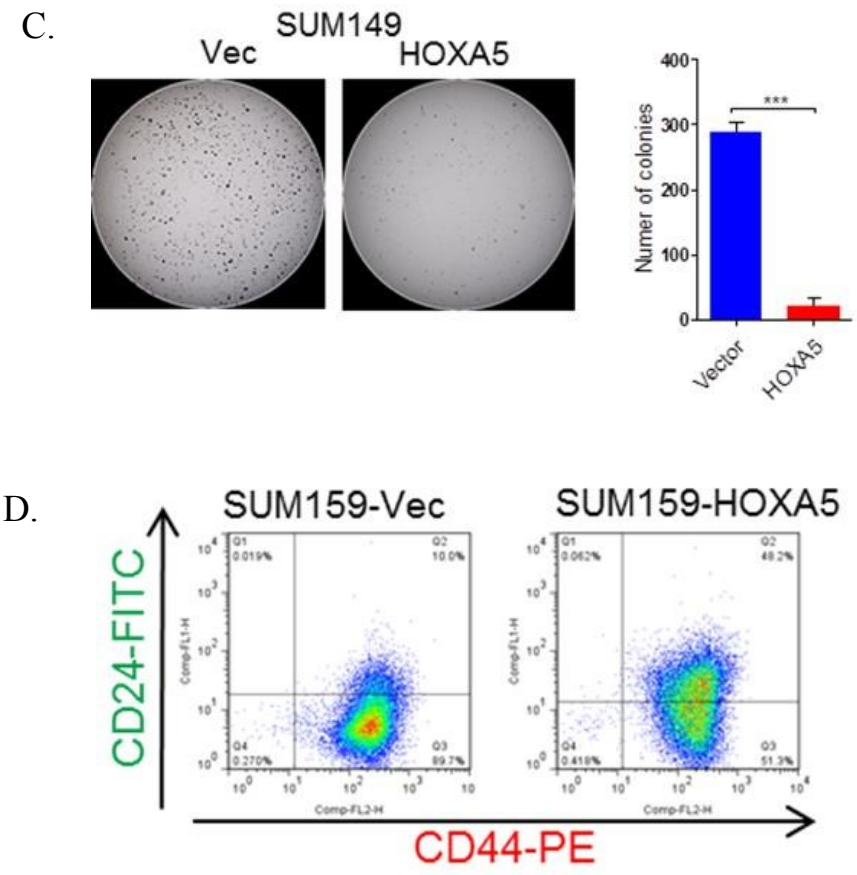
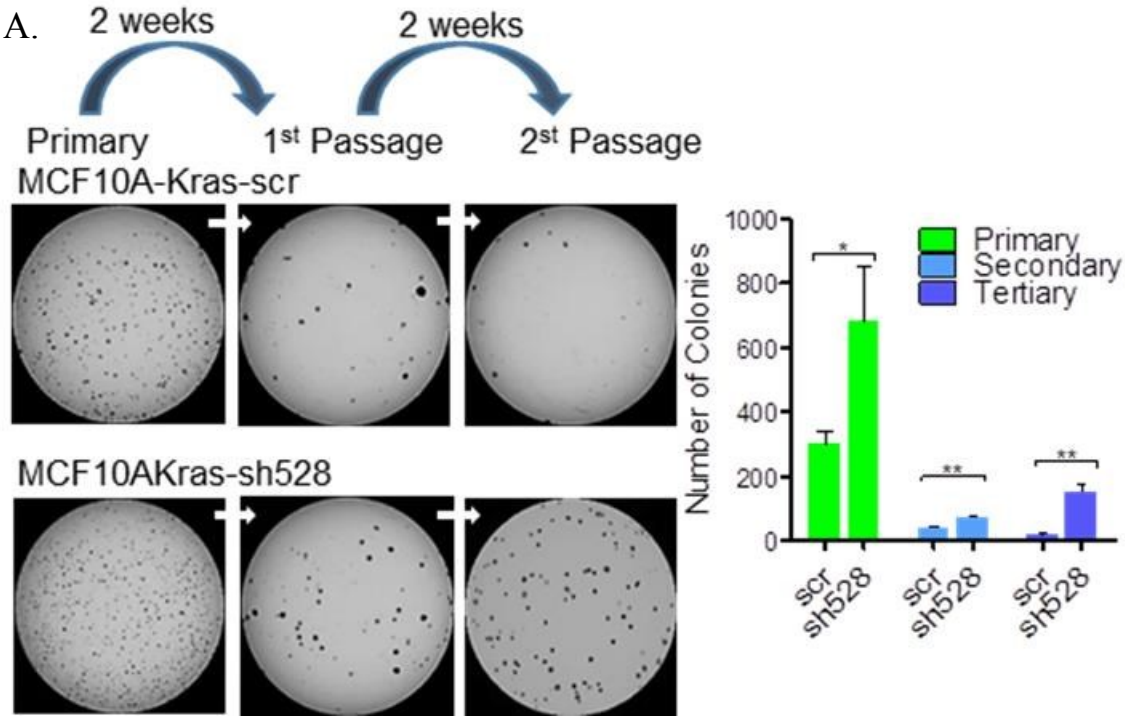


Figure 3.5-3. *HOXA5 inhibits cell stemness in SUM149. A. Flow cytometry analysis of CD24 (x-axis) against CD44 (y-axis) in SUM149-vector and SUM149-HOXA5 cells. B. Quantitation is shown in the bar graphs. C. Representative images and quantitative analysis of mammosphere formation by SUM149-vector and SUM149-HOXA5 cells. D. Flow cytometry analysis of CD24 (x-axis) against CD44 (y-axis) in SUM159-vector and SUM159-HOXA5 cells.*

Our data suggested an inverse correlation between the expression of HOXA5/CD24 and mammosphere formation. Therefore, we sought to determine if HOXA5 reduces self-renewal capacity in breast cancer cells. Serial passaging (Dontu et al 2003) was performed to determine the self-renewal ability of the MCF10A-Kras-scr and -sh528 cells. Depleting HOXA5 in MCF10A-Kras cells markedly enhanced

efficiency of mammosphere formation over the two passages (Figure 3.5-4A); in contrast, SUM149-HOXA5 cells formed few colonies in the first passage that could not be propagated further (Figure 3.5-4B).



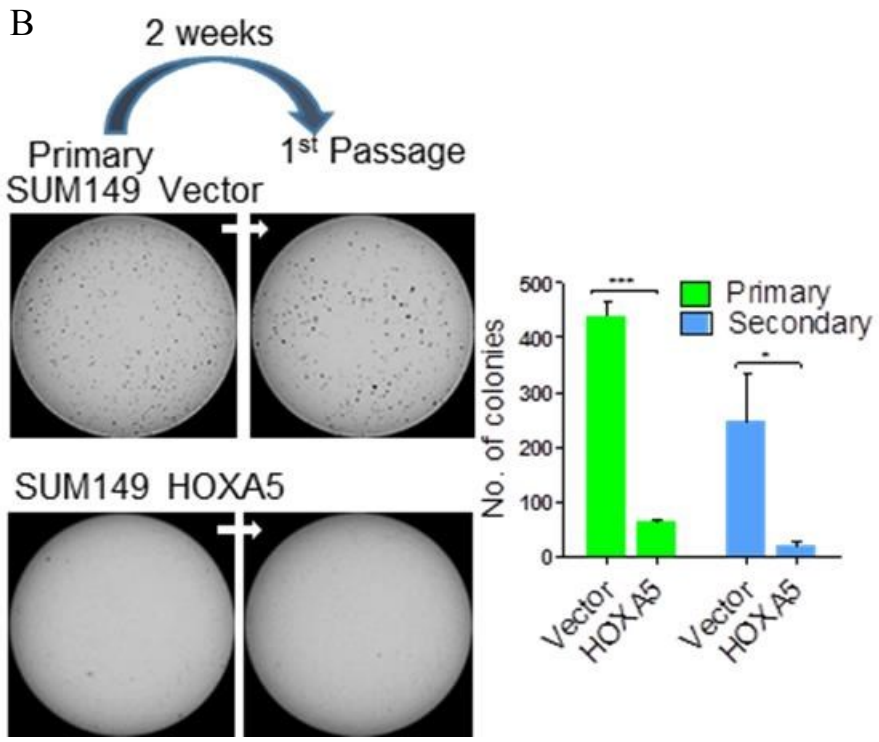
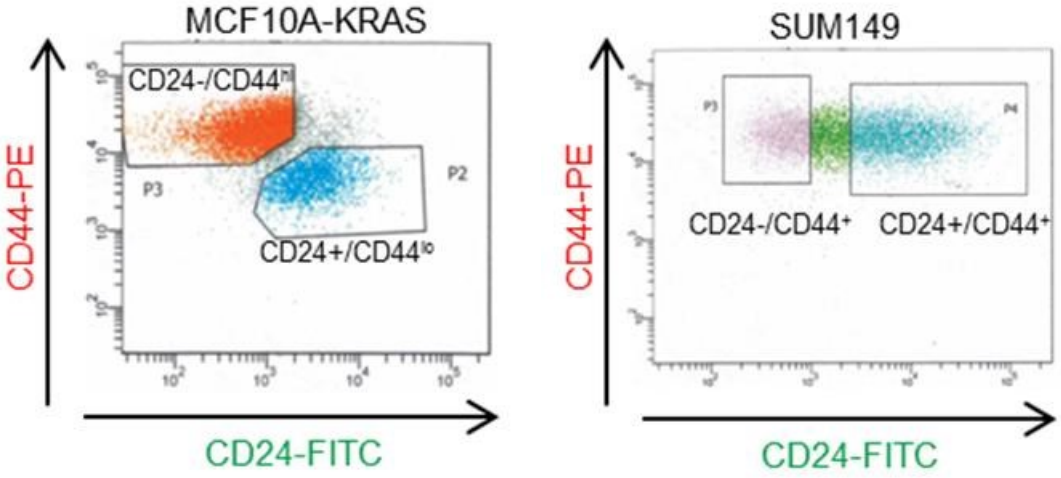


Figure 3.5-4. *Loss of HOXA5 promotes cell renewal capacity in breast cancer cells. Serial passage of mammospheres of: A. MCF10-Kras-scr and HOXA5 depleted -sh528 cells B. SUM-149-vector and -HOXA5 cells; each passage spans two weeks. Quantitative analysis of mammospheres formed in each passage is shown in the bar graphs.*

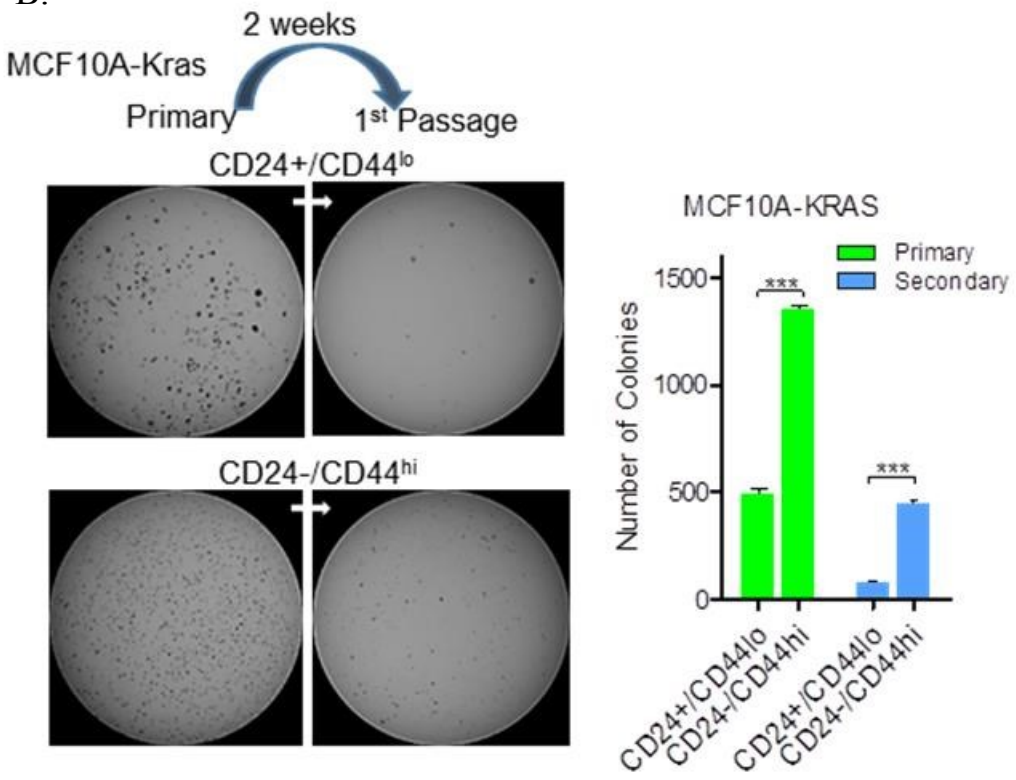
We further sorted breast cancer cell lines, MCF10A-KRAS and SUM149, with CD24 and CD44 cell surface markers and tested if $CD24^{-/lo}$ population is capable of forming more mammospheres in vitro. As predicted, sorted $CD24^{+}/CD44^{lo}$ population Figure 3.5-5A) from MCF10A-KRAS cells formed significantly less number of mammospheres compared to the $CD24^{-}/CD44^{hi}$ counterpart (Figure 3.5-5B) and exhibited lesser self-renewal capacity in forming mammosphere on passage. On the other hand, $CD24^{-}/CD44^{+}$ population isolated from SUM149 formed significantly more

mammospheres than $CD24^{+}/CD44^{+}$ population (Figure 3.5-5C). These findings suggest that expression of CD24 determines the intrinsic capability of mammosphere formation in breast cancer cell lines.

A.



B.



C.

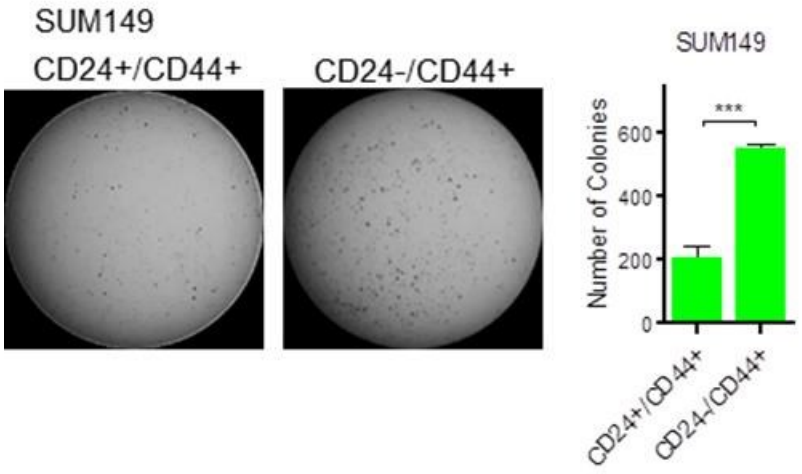


Figure 3.5-5. *CD24-/lo identifies mammosphere forming cells in breast cancer. A. Flow cytometry cell sorting of $CD24^{+}/CD44^{lo}$ and $CD24^{-}/CD44^{hi}$ population from MCF10A-Kras cells or $CD24^{+}/CD44^{+}$ and $CD24^{-}/CD44^{+}$ populations from SUM149 cells B. Serial passage of mammospheres of $CD24^{+}/CD44^{lo}$ and $CD24^{-}/CD44^{hi}$ cells isolated from MCF10A-Kras cells; each passage spans two weeks. Quantitative analysis of mammospheres formed in each passage is shown in the bar graphs. C. Representative image and quantitative analysis of mammosphere formation by $CD24^{+}/CD44^{+}$ and $CD24^{-}/CD44^{+}$ populations isolated from SUM149 cells.*

Supporting these observations, CD24 and CDH1 mRNA expression was lost in MCF10A-Kras-sh528 cells (Figure 3.5-6A). CDH1 protein was undetectable in HOXA5-silenced, MCF10A-Kras-sh528 cells. Conversely, CDH1 protein levels increased by 2-fold in SUM149-HOXA5 cells (Figure 3.5-6B). Collectively, these findings strongly suggested that HOXA5 is functionally necessary, but not sufficient, to reverse stemness in breast cancer cells.

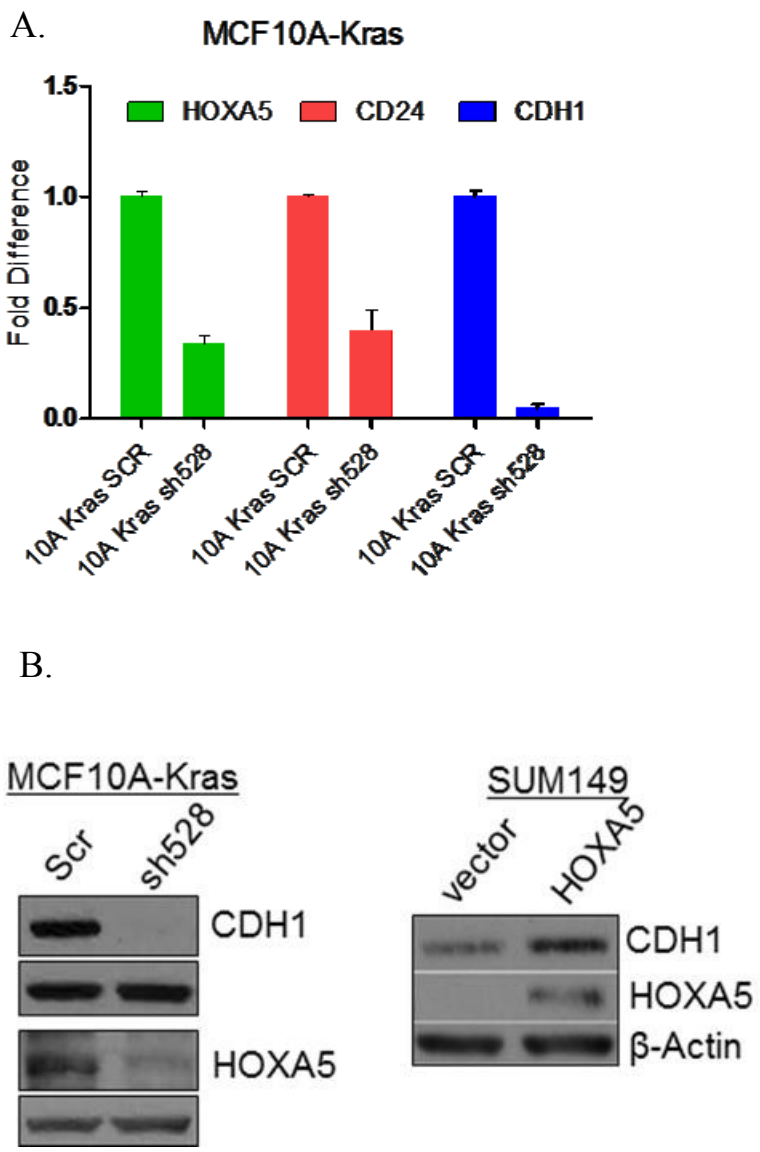
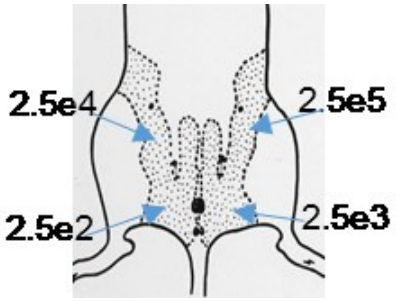


Figure 3.5-6. *HOXA5 expression is associated with CDH1 and CD24 expression in MCF10A-Kras and SUM149 cells. A. Quantitative RT-PCR of HOXA5, CDH1 and CD24 expression level in MCF10A-Kras-scr and MCF10A-Kras-sh528 cells. B. Western Blot analysis of HOXA5 and CDH1 protein in MCF10A-Kras-scr, -sh528, SUM149-vec and -HOXA5 cells. β-actin serves as the loading control.*

3.6 HOXA5 delays tumor initiation and controls tumor differentiation.

Isolated $CD24^{-/low}/CD44^{+}$ population from breast cancer cells or solid tumors has been reported to contain tumor-initiating or cancer stem cells that are more tumorigenic than their counterparts in vivo. Since ectopic expression of HOXA5 leads to increase in $CD24^{+}$ population in SUM149 cells, we evaluated the stemness of SUM149 tumor cells in vivo by performing limiting dilution transplantation experiments using orthotopically injected xenografts. Serially diluted SUM149-HOXA5 cells were injected into humanized mammary fat pads (hmfp) (schema in Figure 3.6-1A). The frequency of palpable tumors was first assessed at the beginning of the fourth-week. A nearly 17-fold reduction in the frequency of stem cells (Figure 3.6-1B) and reduction in tumor volume (Figure 3.6-1C) was observed in the SUM149-HOXA5 cells compared to SUM149-vec control cells. These in vivo findings confirmed our observations in the mammosphere formation assay of SUM149 in culture that HOXA5 has the capacity to induce differentiation, thereby possibly reducing the number of stem cells in the tumorigenic xenograft.

A.



B.

Cell Numbers	SUM149 Vector	SUM149 HOXA5
2.5 x 10^5	7/7	6/7
2.5 x 10^4	7/7	2/7
2.5 x 10^3	2/7	0/7
2.5 x 10^2	0/7	0/7
Stem Cell Frequency	**1 in 6,446**	**1 in 114,319**

C.

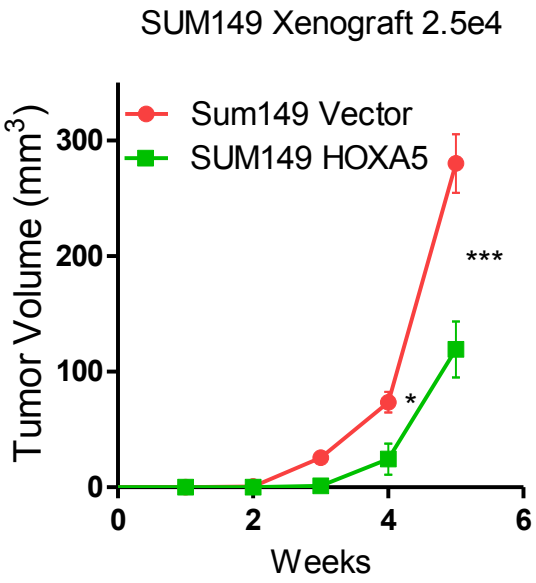


Figure 3.6-1. *HOXA5 delays tumor initiation and tumor growth of SUM149 cells in mice. A. Sites of orthotopic injection of serially diluted SUM149 cells in the 4th and 5th mammary fat pads of NOG mice. B. Table shows xenograft incidence at 3 weeks after orthotopic injection of SUM149-vector or SUM149-HOXA5 cells (n=7 mice for each dilution). Stem cell frequency was evaluated by L-Calc software (with 95% confidence). C. Tumor growth of SUM149 vector vs HOXA5 xenografts following injection of 2.5 x 10^4 cells (n=7 mice). Tumor volumes at each time point are means with SEM, and statistical differences in tumor volumes was analyzed by performing two-tailed student's t test.*

We confirmed these findings in an additional model system. In contrast to HOXA5 over-expressing SUM149 cells, HOXA5 depleted MCF10A-Kras-sh528 xenografts appeared significantly earlier and grew more rapidly than vector control cells (Figure 3.6-2).

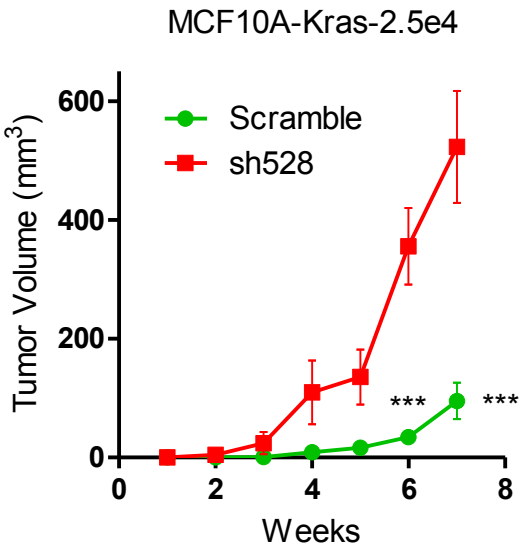


Figure 3.6-2. *Loss of HOXA5 promotes tumor growth of MCF10A-Kras cells in vivo. A. Tumor growth of MCF10A-KRAS-scr vs -sh528 xenografts over 7 weeks following injection of 2.5 x 10^4 explanted cells (n=9 mice). Tumor volumes at each time point are means with SEM, and statistical differences in tumor volumes was analyzed by performing two-tailed student's t test.*

Histopathology of the tumors showed a well differentiated histopathology, characterized by the formation of pseudo-ductal structure with stronger counterstaining with eosin (Figure 3.6-3, left panels, low and high magnification). MCF10A-Kras-sh528 tumors, on the other hand, failed to form duct-like structures and tumor cells exhibited a mesenchymal and dedifferentiated morphology (Figure 3.6-3, right panels).

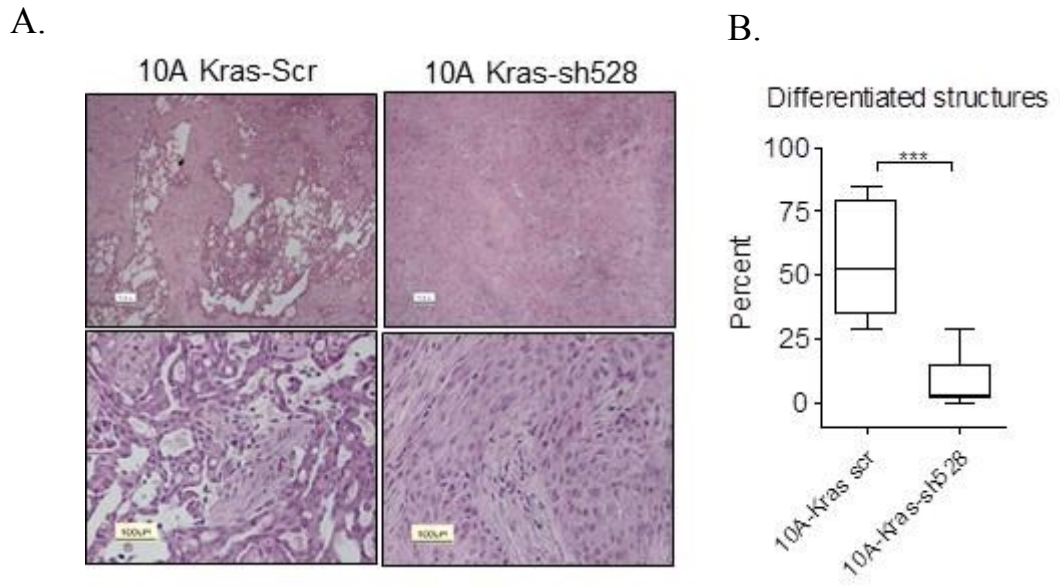


Figure 3.6-3. *Loss of HOXA5 results in dedifferentiation of MCF10A-Kras in vivo. Representative H/E stained histology images of MCF10A-KRAS-scr and HOXA5-depleted MCF10A-KRAS-sh528 xenografts. Quantitative analysis of pseudo-ductal structure formed in MCF10A-KRAS-scr and HOXA5-depleted MCF10A-KRAS-sh528 xenografts.*

Further characterization of the xenografts by IHC showed positive nuclear staining for HOXA5, adhesion junction staining for CDH1 and cytoplasmic staining for CK18 in cells surrounding the pseudo-glandular structures in MCF10A-Kras xenografts while immunostaining was negative for these markers in MCF10A-Kras-sh528 tumors (Figure 3.6-4). The morphology of the cells and loss of luminal epithelial markers provided further evidence that the loss of HOXA5 resulted in the formation of more aggressive and dedifferentiated tumors in vivo.

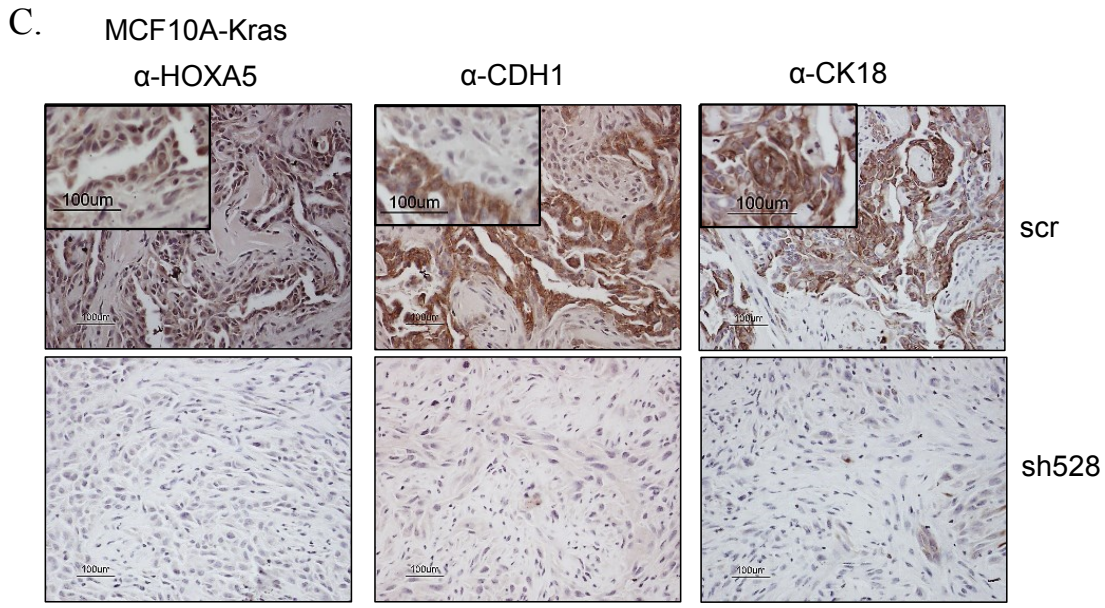


Figure 3.6-4. *IHC characterization of MCF10A-Kras xenografts. Representative images are shown of IHC for HOXA5, CDH1 and CD18 in MCF10A-KRAS-scr and -sh528 tumors. Note nuclear staining for HOXA5, and cell surface staining for CDH1 and cytokeratin, CD18.*

These findings led us to examine the correlations between HOXA5 and the clinical factors in breast cancer. Meta-analysis revealed an inverse correlation between HOXA5 and grade of breast cancer in both the publicly available NKI and Desmedt data sets (Figure 3.6-5A and B). Furthermore, KM-Plot analysis also showed that HOXA5 expression correlated with better relapse free outcome in breast cancer patient (Figure 3.6-4).

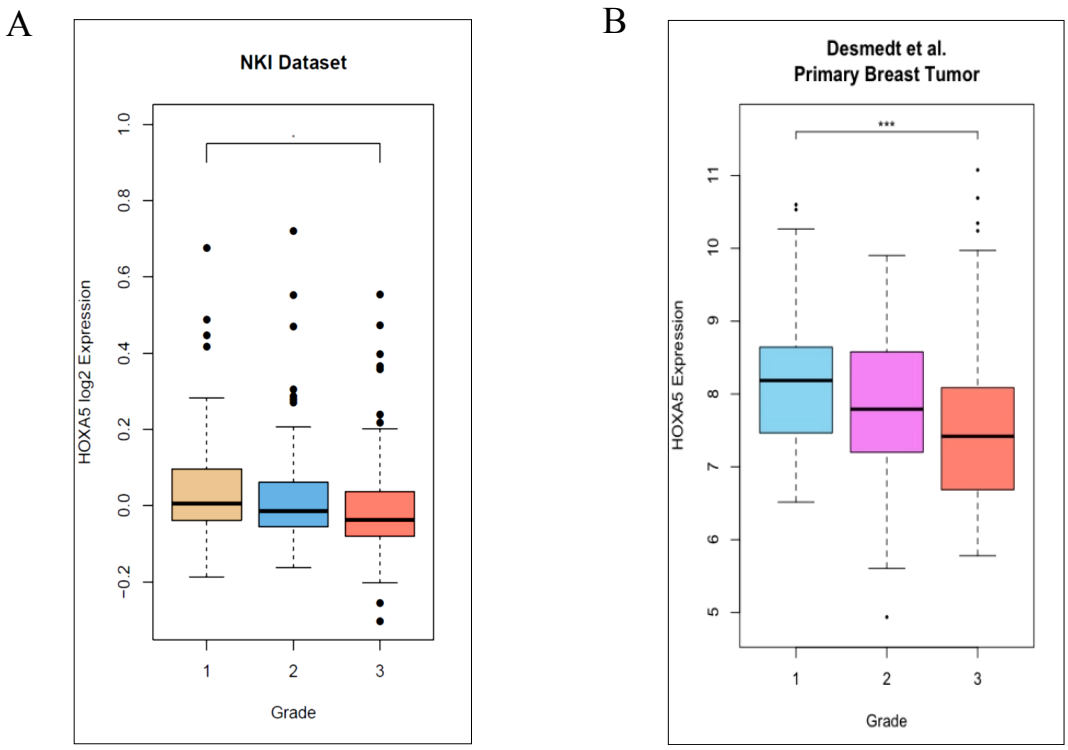


Figure 3.6-5. *HOXA5 inversely correlates with pathological grades of breast cancer. In silico analysis of correlation between HOXA5 transcript level and pathological grades of primary breast tumors from: A. NKI and B. Desmedt datasets.*

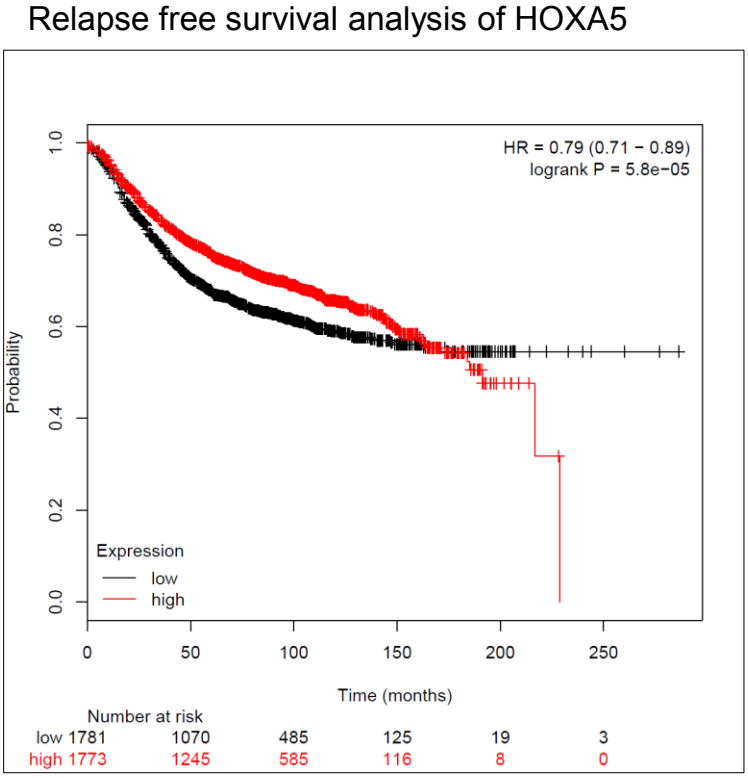
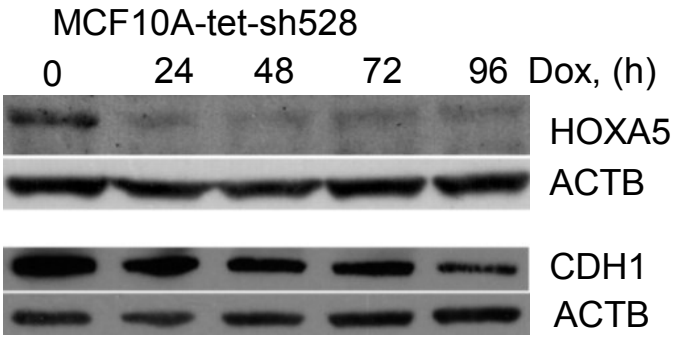


Figure 3.6-6. *HOXA5 predicts better prognosis in breast cancer patients. Relapse free survival curve of HOXA5 was derived with a web-based tool-Kaplan Meier plotter using median as the cut off (High>median, Low<median).*

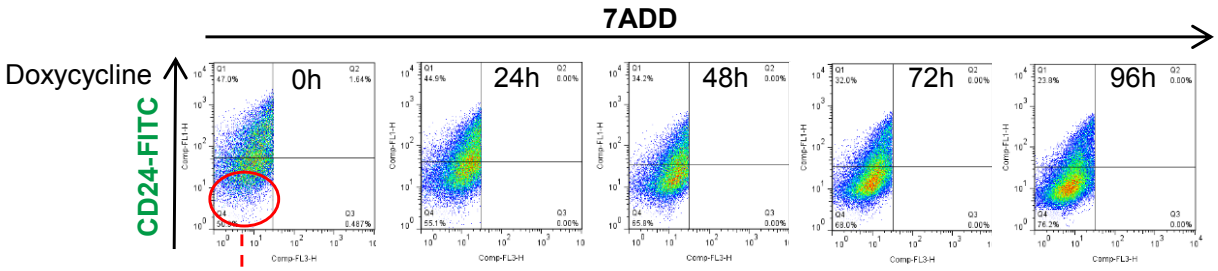
3.7 Differentiation of the $CD24^-/CD44^+$ population by retinal is partially mediated by HOXA5.

Our findings clearly highlight HOXA5's function in suppressing EMT characteristics, but a possible dynamic role for HOXA5 during cell fate transition is yet to be addressed. Pathway analysis of the differentially expressed genes in HOXA5-depleted MCF10A cells revealed changes in genes involved in retinoic acid biosynthesis, such as ALDH1A3 and DHRS3. Retinoic acid signaling and HOX genes have long been implicated in tissue differentiation and development (Strickland and Mahdavi 1978, Wang et al 2005). Moreover, previous publications from others and us have provided evidence that retinoic acid receptors are direct transcriptional regulators of HOXA5 (Chen et al 2007). We developed an inducible HOXA5 knockdown model in MCF10A cells that would allow us to study the dynamics of this transition. Doxycycline-induced loss of HOXA5 in MCF10A-tet-sh528 resulted in a sequential loss of CDH1 protein level and $CD24^+$ population over time (Figure 3.7-1A and B) and loss of expression of CD24 and CDH1 mRNA, an effect not observed in MCF10A-vec cells treated with doxycycline (Figure 3.7-1D). As seen in Figure 3.7-1A and in the Western blot in Figure 3.7-1B, reduction of HOXA5 expression was observed as early as 24h, followed by reduction in CDH1 and CD24 protein levels. $CD24^-$ population in MCF10A-tet-sh528 cells was isolated by flow-cytometry cell sorting, and the cells were cultured for one week in vitro. Cells were treated with retinal to induced differentiation with or without the addition of doxycycline to study the functions of HOXA5 in the differentiation cascade (schema Figure 3.7-1A-C).

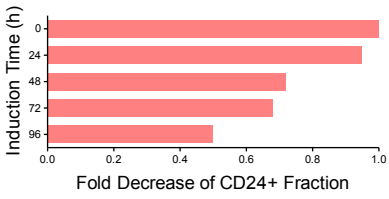
A.



B.

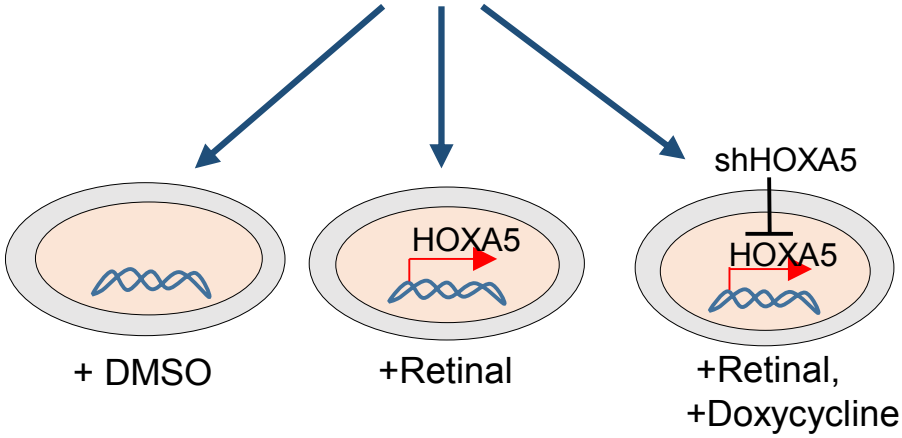


C.



Sorted CD24^{-neg} MCF10A-tet-sh528 cells cultured for 1 week

D.



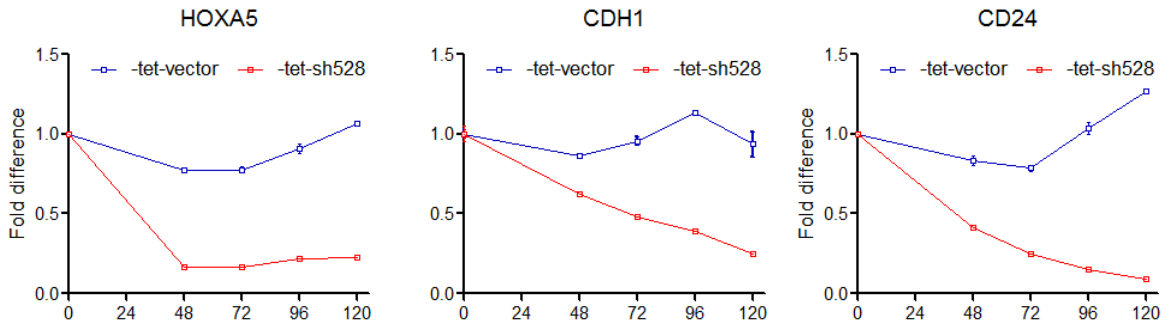
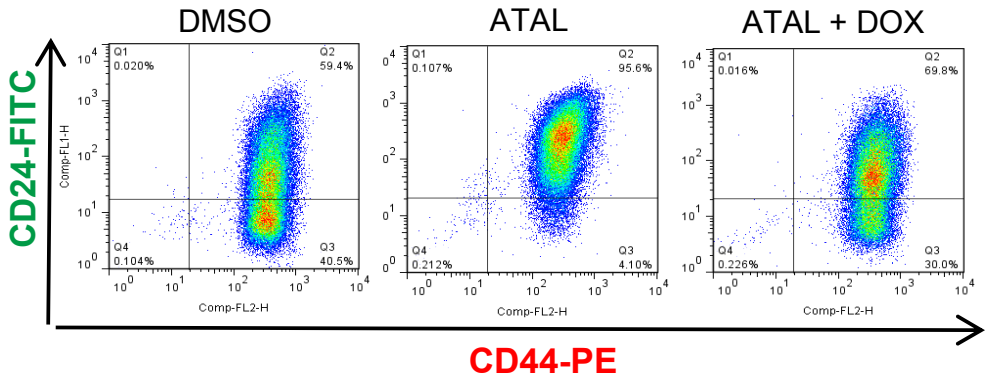


Figure 3.7-1. *CDH1 and CD24 are early downstream targets of HOXA5 in MCF10A cells. A. Western blot analysis of HOXA5 and CDH1 expression in MCF10A-tet-sh528 cells with dox-induced depletion of HOXA5 over 96 hours. β-actin serves as the loading control. B. Flow cytometry analysis of CD24 expression in MCF10A-tet-HOXA5 with doxycycline-induced depletion of HOXA5 over 96 hours. Quantitative analyses of* $CD24^{+}/CD44^{+}$ *and* $CD24^{-}/CD44^{+}$ *populations is shown in the bar graph. C. Flow sorted* $CD24^{-}$ *population was treated with DMSO, 1 μM retinal (ATAL) or 1 μM ATAL + 100 nM doxycycline (DOX) and cultured for 1 week. D. Quantitative RT-PCR analysis of HOXA5, CDH1 and CD24 expression over 120 hours doxycycline treatment in MCF10A-tet-vector and MCF10A-tet-sh528 cells.*

Compared to the number of $CD24^{+/}CD44^{+}$ cells in DMSO-treated cells, treatment with retinal (1uM) resulted in a striking enrichment (40% to 96%) of this population. Concurrent repression of HOXA5 with doxycycline along with the HOXA5 inducer, retinal, resulted in partial abrogation of retinal-mediated increase in $CD24^{+/}CD44^{+}$ population (96% to 70%)(Figure 3.7-2A). Bar graphs show quantitation of $CD24^{+/}CD44^{+}$, and of $CD24^{-}/CD44^{+}$ populations (Figure 3.7-2B).

A.



B.

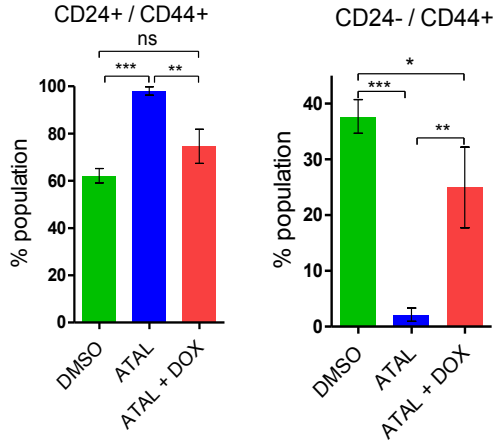
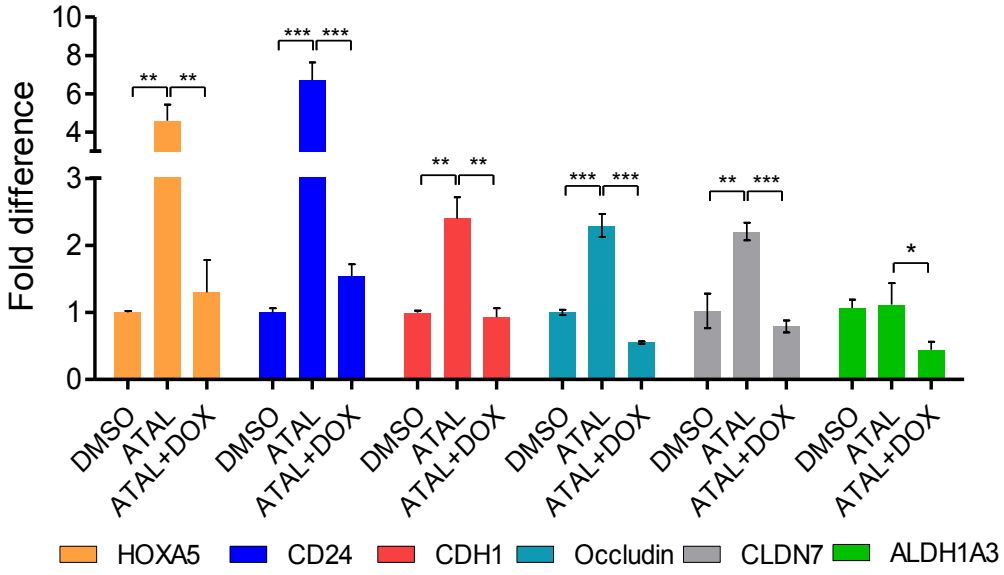


Figure 3.7-2. *HOXA5 mediates upregulation of CD24 in MCF10A cells by retinal. A. Flow cytometry analysis of CD24 (x-axis) and CD44 (y-axis) in the cells treated as shown in (3.7-1C). B. Quantitative analyses of $CD24^+/CD44^+$ and $CD24^-/CD44^+$ populations is shown in the bar graph*

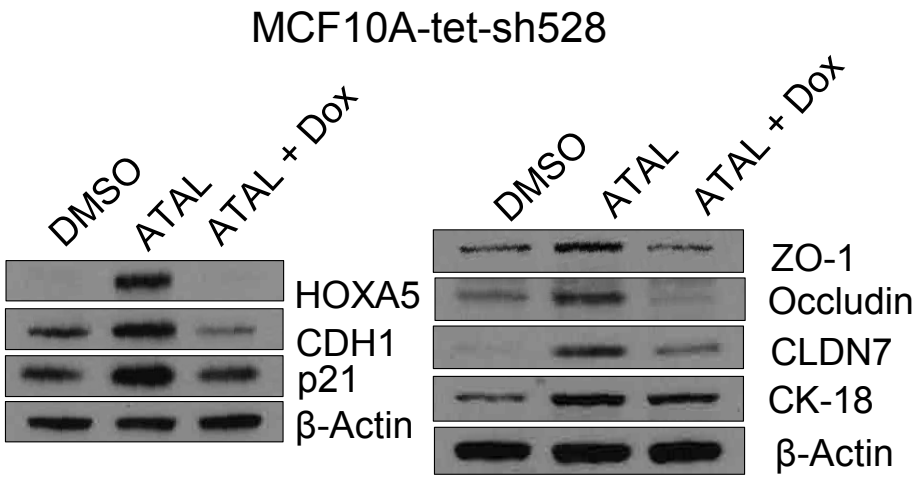
mRNA expression of CD24, CDH1, Occludin and CLDN7, but not ALDH1A3 was induced by retinal, and all the genes were repressed when HOXA5 was simultaneously silenced in MCF10A (3.7-3A). Differentiation induced by retinal is further characterized by the elevation of junctional and structural proteins which expression is epithelial specific. Western blot analysis of CDH1, ZO-1, Occludin, CLDN7, CK18 and cell cycle arrest protein, p21, showed elevated expression upon retinal treatment which was reversed by depleting HOXA5 (Figure 3.7-3B). Taken together, we concluded that retinal-induced cell differentiation in MCF10A cells is, in part, driven by HOXA5. Further substantiation was sought through additional assays of stem cell differentiation in MCF10A-tet-sh528 cells. Retinal inhibited mammosphere

formation while concurrent knockdown of HOXA5 in the presence of retinal significantly restored the ability of the cells to form mammospheres (3.7-3C). These findings strongly suggested that HOXA5 is one of the key mediators of retinal-induced differentiation in MCF10A cells.

A.



B.



C.

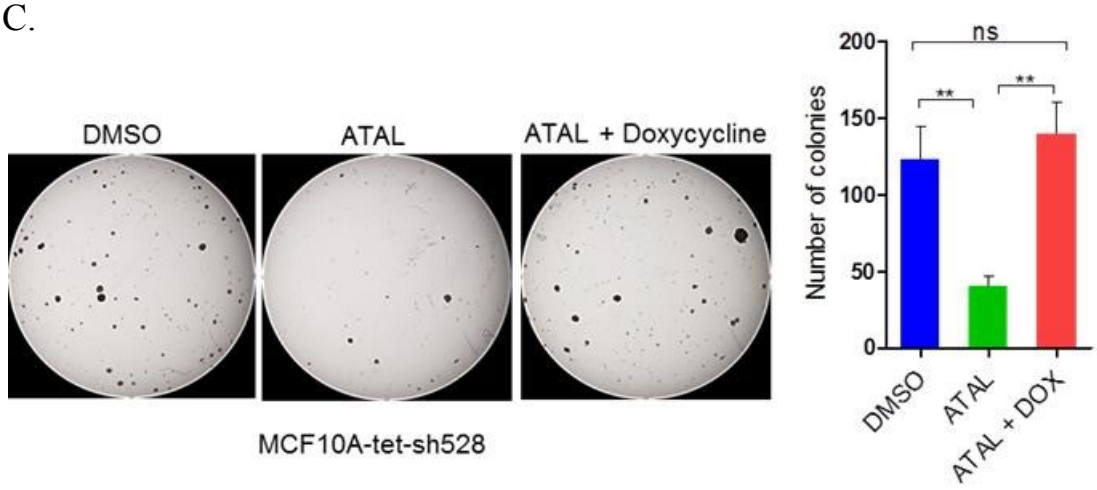


Figure 3.7-3. *HOXA5 mediates differentiation of MCF10A cells by retinal.* A. *Quantitative RT-PCR analysis of HOXA5, CD24, CDH1, Occludin, CLDN7 and ALDH1A3 in MCF10A-te-sh528 cells treated as indicated in (3.7-1C)* B. *Western blot analysis of HOXA5, CDH1, p21, ZO-1, Occludin, CLDN7 and CK18 in the cells treated as shown in (3.7-1C). β-actin serves as the loading control.* C. *Representative images and quantitative analysis of mammosphere formation in MCF10A-tet-sh528 cells treated with DMSO, ATAL, ATAL + DOX for one week.*

3.8 HOXA5 is a direct transcriptional regulator of CDH1 and CD24.

How does HOXA5 regulate the multiple traits typical of luminal epithelial cells in the mammary gland? We addressed this question for CDH1 and CD24 which showed strong correlation with HOXA5 expression. We performed promoter-linked luciferase assays for CDH1 and CD24 to investigate whether expression of HOXA5 trans-activate the promoters of CDH1 and CD24. Transient expression of HOXA5 along with the luciferase constructs in 293T cells resulted in a significant induction of luciferase activity driven by the promoters of both CDH1 (Figure 3.8-1A) and CD24 (Figure 3.7-1B). The homeodomain was necessary for this activity in both promoters, since

luciferase activity was significantly abrogated when 293T cells were transfected with HOXA5 (Del HD) instead of wild type HOXA5 (Figure 3.8-1A and 3.8-1B).

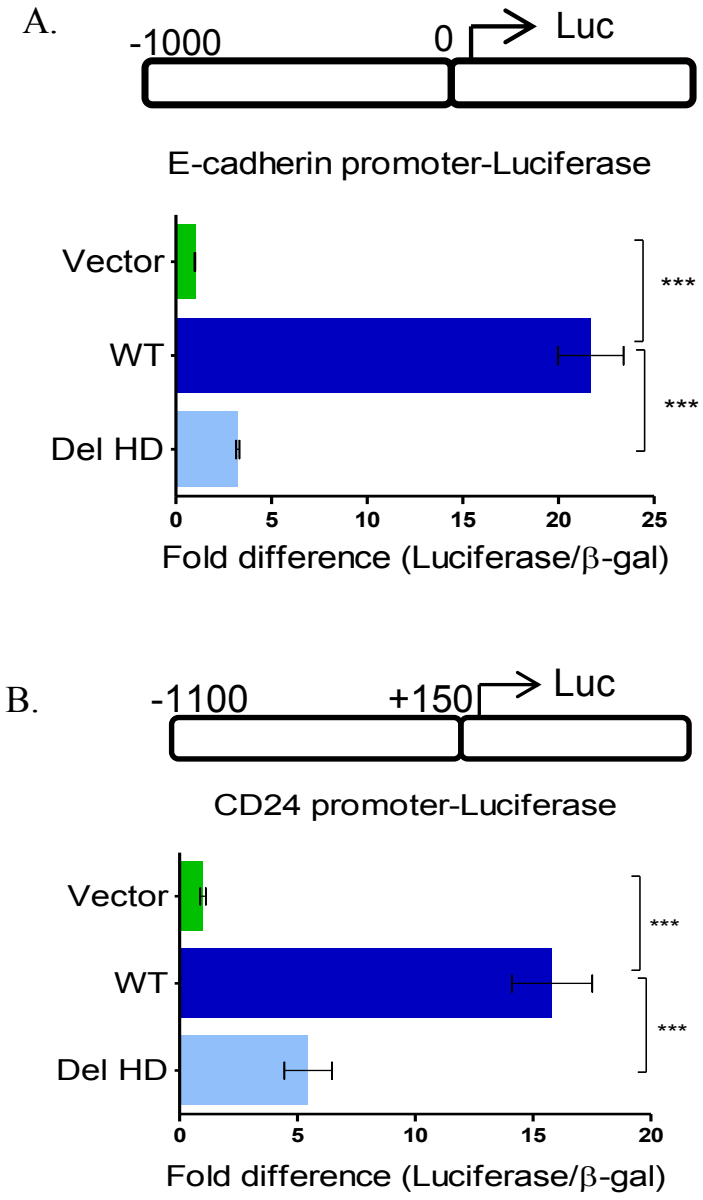
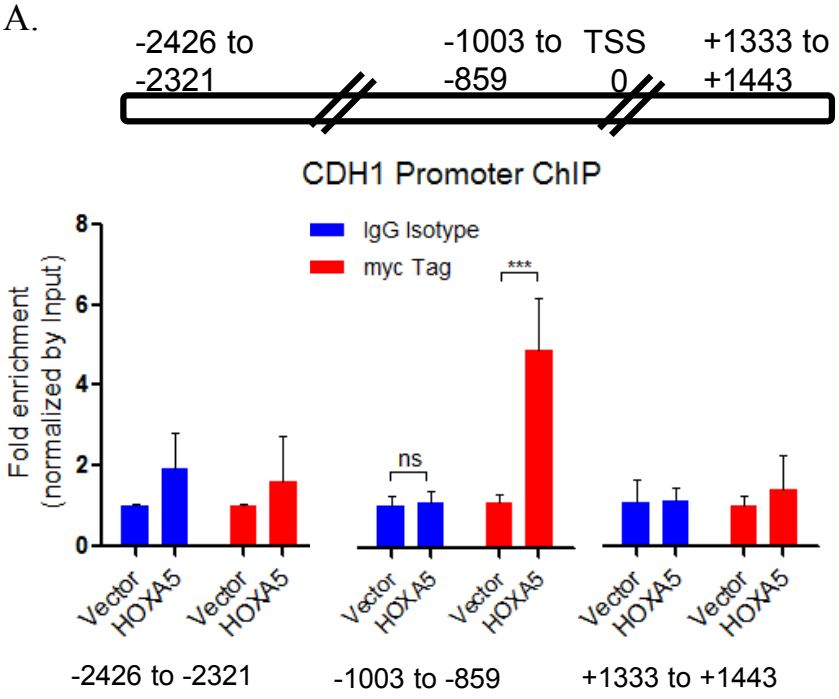


Figure 3.8-1. *HOXA5 increases promoter-luciferase activity of CDH1 and CD24. Promoter linked luciferase activity was analyzed in 293T cells, 24 h after cotransfection of* ***A.*** *a luciferase construct containing the promoter sequence (-1000 bp) of E-cadherin or vector control, or* ***B.*** *a luciferase construct containing the CD24 promoter (-1100 bp to +150bp) and vector control, each with wild type (WT) or homeodomain truncated*

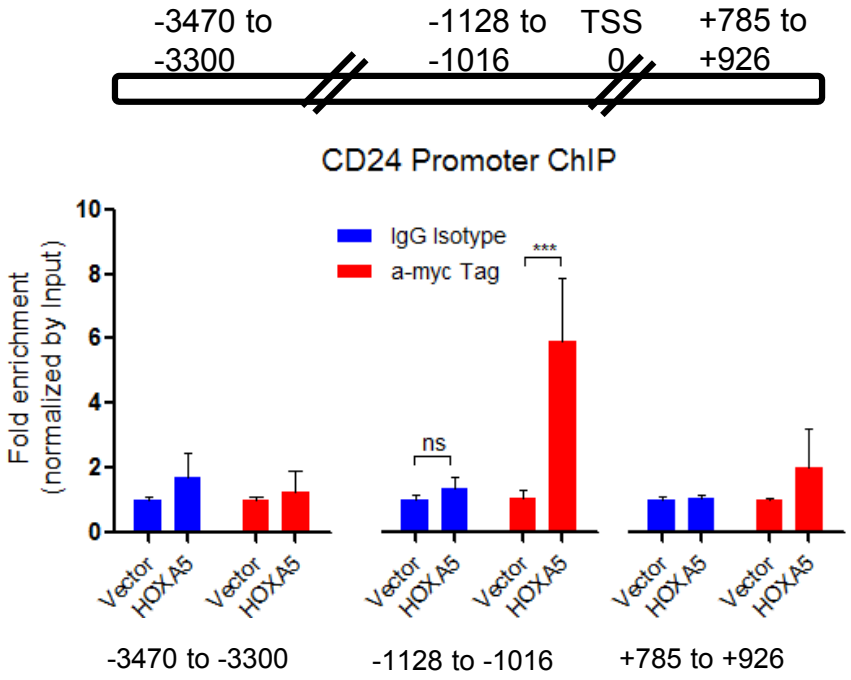
(ΔHD) HOXA5, and β-galactosidase plasmid to control for transfection efficiency shows that HOXA5 upregulated luciferase activity of both promoters, which was significantly attenuated when the HD was truncated.

We further performed ChIP assays on MCF10A cells overexpressing HOXA5 (myc-tag) to determine the occupancy of HOXA5 at a putative HOX-binding-site (TAAT) located in each of the gene promoters. In both cases, we observed statistically significant enrichment of HOXA5 at their respective sites in chromatin, compared to the vector controls, confirming HOXA5 occupancy at the region. In parallel, no significant ChIP enrichment by HOXA5 was detected at two distanced sites which are approximately 2000bp upstream or downstream of these regions of promoters of CDH1 or CD24 (Figure 3.8-2A and 3.8-2B). In sum, these results suggested that HOXA5 is a transcriptional regulator of two luminal-specific genes, CDH1 and CD24. Next, we sought confirmation in cells with induced expression of the endogenous HOXA5 gene to determine if HOXA5 induced by retinal treatment in MCF10A cells could bind to the CD24 promoter. ChIP analysis showed that HOXA5, induced by retinal, bound to the promoter region of CD24 in MCF10A cells. Moreover, when HOXA5 was depleted by addition of doxycycline, its occupancy at the same region was significantly reduced. The enrichment, however, was not detected at the distant regions which are approximately 2000 bp upstream or downstream from this site (Figure 3.8-2C). This finding strengthens the transcriptional role of HOXA5 in retinal mediated induction of CD24 expression.

A.


-2426 to -2321
-1003 to -859
TSS 0
+1333 to +1443
CDH1 Promoter ChIP
IgG Isotype
myc Tag
***
ns
Fold enrichment (normalized by Input)
0
2
4
6
8
Vector
HOXA5
-2426 to -2321
-1003 to -859
+1333 to +1443

B.


-3470 to -3300
-1128 to -1016
TSS 0
+785 to +926
CD24 Promoter ChIP
IgG Isotype
a-myc Tag
***
ns
Fold enrichment (normalized by Input)
0
2
4
6
8
10
Vector
HOXA5
-3470 to -3300
-1128 to -1016
+785 to +926

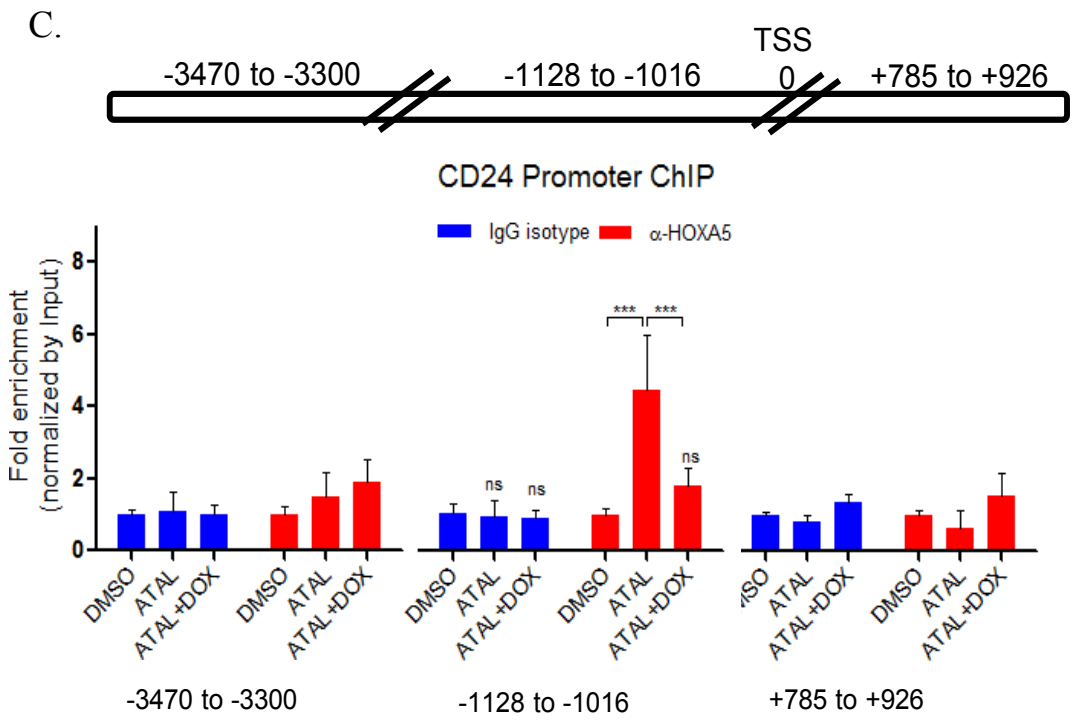


Figure 3.8-2. *HOXA5 occupancy at the promoters of CDH1 and CD24.* A. *ChIP analysis of the E-cadherin promoter was performed in vector or myc-tagged-HOXA5 expressing MCF10A-sh528 cells using α-myc antibody and isotype IgG control. PCR shows enrichment of occupancy of HOXA5 at the -1003 bp to -859 bp region of E-cadherin containing putative TAAT HOX-binding site upstream of its transcription start site but not at the upstream (-2426 bp to -2421 bp) and downstream (+1333 to +1443) regions. There was no enrichment in the IgG ChIP.* B. *ChIP analysis of CD24 promoter was performed in vector or myc-tagged HOXA5 expressing MCF10A-sh528 cells using α-myc antibody and isotype IgG control. Bar graph showing enrichment of occupancy of HOXA5 at the -1128 bp to -1016 bp region of CD24, containing a putative TAAT HOX- binding site upstream of CD24 transcription start site but not at the upstream (-3470 bp to -3300 bp) and downstream (+785 to +926) regions. There was no enrichment in the IgG ChIP.* C. *ChIP analysis of CD24 promoter in DMSO, ATAL or ATAL+DOX treated MCF10A cells using α-HOXA5 antibody or isotype IgG control. Bar graph shows enrichment of occupancy of HOXA5, but not when HOXA5 was depleted, at the -1128bp to -1016bp region of CD24 promoter. There was no*

enrichment at the upstream (-3470 bp to -3300 bp) and downstream (+785 to +926) region.

4 Discussion

Our investigation has revealed a role for HOXA5 in regulating cell differentiation, cell specification and tumor initiation in breast cancer, and provided mechanistic insight into how HOXA5 maintains homeostasis in mammary epithelial cells. Microarray analysis of human mammary epithelial cells identified E-cadherin (CDH1) and CD24 as key molecules whose expression is lost as a consequence of loss of HOXA5. We present evidence that HOXA5 is a direct transcriptional regulator of both CDH1 and CD24. Further, HOXA5 loss leads to reduced CDH1 expression and subsequent development of hallmarks of epithelial mesenchymal transition, invasion and migration, while reduced CD24 expression and consequent enrichment of $CD24^{-}$ /$CD44^{+}$ population leads to an expansion of breast cancer stem cells and rapid tumor growth.

Depletion of HOXA5 in MCF10A cells resulted in acquisition of cell motility, and invasiveness and an ability to form mammospheres, suggestive of a tumor suppressive role for HOXA5 in cancer progression. Loss of (CDH1), tight junction protein (CLDN1), desmosome proteins (DSG3) and luminal fate marker, CD24 and down-regulation of retinoic acid metabolic enzymes, ALDH1A3 and DHRS3 were all, collectively, indicators of the disruption of differentiation signals. The findings that HOXA5 depletion results in loss of structural proteins of the cells are in agreement with the findings that HOXA5 can transcriptionally regulate multiple members of the actin cytoskeleton in non-small cell lung carcinomas (Wang et al 2015) and can impede cell

migration and invasion in culture (Korangath et al 2015, Wang et al 2015). Our findings uncover a possible link between HOXA5 and epithelial integrity, which when disrupted, leads to the cancer development.

Although members of HOX5 genes have been implicated in cellular differentiation, most of the evidence was gathered from non-mammalian models (Henderson and Andrew 2000). The size of HOXA5 null mice was found to be smaller than heterozygous littermates, and many died within a week of birth due to respiratory tract defects (Aubin et al 1997). Survivors showed delayed and abnormal mammary development during pregnancy, leading to production of very little or no milk, but did not develop mammary tumors. In the mouse, the investigators showed that HOXA5 loss in the stroma, not the epithelium, was responsible for abnormalities in development of the mammary gland (Garin et al 2006). In recently developed HOXA5-cre models (Berube-Simard and Jeannotte 2014), skeletal and other abnormalities were detected, but no mammary tumors were reported, indicating that loss of HOXA5 alone is insufficient for tumorigenesis. Technical limitations have hindered the exploration of the mechanism of regulation of mammary cell differentiation by HOXA5. Our studies were facilitated by generating MCF10A mammary epithelial cells with Tet-inducible shHOXA5. Using these cells and other models, our data is supportive of a role for HOXA5 in specifying cell fate brought about by up-regulating critical cell surface proteins that maintain the structural integrity of epithelial cells. Our studies, using promoter-linked luciferase assays and ChIP analysis showed that the expression of CDH1 is transcriptionally regulated by HOXA5 in MCF10A cells. Expression of occludin and claudins was also reduced in cells stably depleted of HOXA5. Like

occludin, CLDN1 and CLDN7 are important components of tight junctions, a barrier expressed primarily in differentiating ductal cells. However, their expression is consistently lost in many breast carcinomas and breast cancer cell lines (Kominsky et al 2003, Kramer et al 2000, Martin et al 2010, Swisshelm et al 1999, Tokes et al 2005). Hence, these findings provide evidence for yet another of the many complex functions of HOXA5- that of regulating structural integrity of epithelial cells and thereby, likely, cellular patterning and development.

We show that silencing of HOXA5 in MCF10A cells leads to the reduction in expression of CD24. Based on the loss of a differentiated cell marker CD24, we predicted that HOXA5 depleted MCF10A will convert to more primitive and basal phenotypes. Using a combination of knockdown of HOXA5 in MCF10A-Kras, and overexpression in SUM149 or SUM159 breast cancer cell lines, we have shown that HOXA5 reduces CSC population defined by expression of CD24/CD44 markers. Increase of the more differentiated $CD24^{+}$ population, reduction of self-renewal capacity accompanied by elevation of CDH1 expression provides strong evidence that HOXA5 impedes stemness of breast cancer cells. Our results support the theory that breast cancer cells might acquire enhanced ability to form tumors when they regress to a more undifferentiated state (Keller et al 2010). In fact, tumor initiating cells were generated through dedifferentiation of epithelial cells in a murine intestinal model (Schwitalla et al 2013). Our results support several tenets of this reversion theory: 1) Expression of HOXA5 markedly reduced the number of tumor initiating cells in limiting dilution assays performed in immunodeficient mice. Also, depletion of HOXA5 in MCF10A-Kras cell line resulted in the formation of histopathologically

dedifferentiated xenografts, an observation that is in agreement with the inverse correlation between HOXA5 expression and pathological grades of breast cancer and poorer survival outcomes. These observations strongly support a differentiating role for HOXA5, which may limit tumor initiating capacity in breast cancer.

It remains unclear what drives the transition from a primitive to the differentiated axis. The inducible HOXA5 knock down model system in MCF10A cells allowed us demonstrate the role HOXA5 in retinal-mediated transition from the CD24- to CD24+ axis, which was characterized by elevated expression of epithelial specific CDH1 and several tight junction proteins. Retinal significantly increased the $CD24^{+}$ population through upregulation of HOXA5 in MCF10A, consequently decreasing their mammosphere-initiating capability. Our findings are consistent with the presumed role of retinoic acid pathway in cellular differentiation (Chute et al 2006, Ghiaur et al 2013, Strickland and Mahdavi 1978) . We further confirmed that HOXA5 is necessary, but not sufficient for the differentiation process. Although cell surface molecules are useful for delineating specific cell lineage, there are not many studies focusing on transcriptional cues for cell state transition. As a regulator of differentiation of the epithelial lineage, we have shown that HOXA5 transcriptionally regulates CDH1 and CD24 through promoter binding. Thus, our findings identify novel molecular pathways to evaluate molecular perturbation during the transition of cell fate in vitro.

This study strongly supports a role for HOXA5-mediated differentiation in mammary ductal development. Our microarray data support the fact that HOXA5 may control tumor environment through regulation of extracellular matrix components, which can contribute to the forming of the differentiated structures. Studies in the future,

could address developmental mechanisms driven by HOXA5 in the presence of environmental cues.

Appendix

Genes differentially expressed between MCF10A-scr and HOXA5 depleted cells that have an FDR < 0.05.

Symbol	Gene	EntrezID	ProbeID	logFC_sh528	FDR_sh528	logFC_sh529	FDR_sh529
CDH1	cadherin 1, type 1, E-cadherin (epithelial)	999	ILMN_177094	-3.1894386	2.08E-08	-1.6383841	7.51E-05
KRT6B	keratin 6B, type II	3854	ILMN_172135	-2.752207	1.77E-06	-2.0366395	0.00013973
S100P	S100 calcium binding protein P	6286	ILMN_180121	-2.4125	1.09E-06	-1.014813	0.00511372
ALDH1A3	aldehyde dehydrogenase 1 family, member A3	220	ILMN_213997	-2.2377303	9.63E-06	-1.2406281	0.0051898
ALDH1A3	aldehyde dehydrogenase 1 family, member A3	220	ILMN_180743	-2.2290761	1.43E-07	-1.319296	0.00011729
CD24	CD24 molecule	1E+08	ILMN_180551	-1.8880121	8.15E-06	-1.0544794	0.00480046
CCND2	cyclin D2	894	ILMN_166708	-1.8598427	1.30E-06	-0.7987185	0.00569265
S100A14	S100 calcium binding protein A14	57402	ILMN_178328	-1.8574984	1.97E-06	-0.8111102	0.00816925
HOXA5	homeobox A5	3202	ILMN_175361	-1.7916322	4.92E-06	-1.2349601	0.00063257
TACSTD2	tumor-associated calcium signal transducer 2	4070	ILMN_173900	-1.7147412	1.97E-06	-0.8786096	0.00324329
CA2	carbonic anhydrase II	760	ILMN_166279	-1.5837588	2.47E-05	-0.683894	0.04884242
PI3	peptidase inhibitor 3, skin-derived	5266	ILMN_169319	-1.4495864	5.43E-06	-1.4924321	5.41E-05
MAPK13	mitogen-activated protein kinase 13	5603	ILMN_174932	-1.4354518	1.04E-05	-0.8355596	0.00425672
MBP	myelin basic protein	4155	ILMN_167266	-1.4321106	5.43E-06	-0.5701811	0.03076032
CLDN1	claudin 1	9076	ILMN_172468	-1.3452025	0.00013689	-0.9029693	0.01050089
HES2	hes family bHLH transcription factor 2	54626	ILMN_209426	-1.2724239	1.82E-05	-1.4834779	5.41E-05
ADAP1	ArfGAP with dual PH domains 1	11033	ILMN_204751	-1.268196	1.98E-05	-0.5687532	0.03365609
DSG3	desmoglein 3	1830	ILMN_165546	-1.2418481	0.00022334	-1.1975338	0.00147958
IRF6	interferon regulatory factor 6	3664	ILMN_172594	-1.114309	0.00085225	-0.7258522	0.04574719
TAGLN	transgelin	6876	ILMN_240093	-1.0875883	0.0009986	-0.9040401	0.01213146
TAGLN	transgelin	6876	ILMN_177866	-1.0239275	0.00192106	-0.8661436	0.01998833
MIR1978	MIR1978		ILMN_331049	-1.0190117	0.00052036	-0.896863	0.00486279
VCL	vinculin	7414	ILMN_241352	-0.9756149	0.00076388	-0.6630518	0.0323792
DHRS3	dehydrogenase/reductase (SDR family) member 3	9249	ILMN_175247	-0.9254824	0.00114832	-0.8914146	0.00502973
KRT16	keratin 16, type I	3868	ILMN_173676	-0.8820226	0.00129175	-1.015344	0.00180092
RNF145	ring finger protein 145	153830	ILMN_171090	-0.8772147	0.00162098	-0.9141301	0.00431284
MYL9	myosin, light chain 9, regulatory	10398	ILMN_177695	-0.8749279	0.00080929	-0.6018149	0.03232725
B4GALT5	UDP-Gal:betaGlcNAc beta 1,4- galactosyltransferase, polypep	9334	ILMN_168582	-0.8587968	0.00251818	-0.7268605	0.02592263
PTP4A1	protein tyrosine phosphatase type IVA, member 1	7803	ILMN_176057	-0.8484173	0.00219948	-1.1228496	0.00119233
CD70	CD70 molecule	970	ILMN_176024	-0.8362173	0.00352638	-1.4961361	0.00017171

Genes differentially expressed between MCF10A-scr and HOXA5 depleted cells that have an FDR < 0.05.

CREG1	cellular repressor of E1A-stimulated genes 1	8804	ILMN_168062	-0.8331741	0.0123252	-0.9240646	0.0202103
YRDC	yrdC N(6)-threonylcarbamoyltransferase domain containing	79693	ILMN_206173	-0.8072973	0.00152451	-0.8466893	0.00389078
GPRC5A	G protein-coupled receptor, class C, group 5, member A	9052	ILMN_168259	-0.8013007	0.0021334	-0.6793256	0.0217221
CYFIP2	cytoplasmic FMR1 interacting protein 2	26999	ILMN_167720	-0.7950999	0.00162763	-0.8520362	0.00367251
KRT16P3	keratin 16 pseudogene 3	644945	ILMN_175074	-0.7446781	0.01113484	-0.7074434	0.04443826
PHLDA2	pleckstrin homology-like domain, family A, member 2	7262	ILMN_167155	-0.7303423	0.00351439	-0.6658025	0.0216497
ST6GAL1	ST6 beta-galactosamide alpha-2,6-sialyltranferase 1	6480	ILMN_175650	-0.7136165	0.00214979	-0.5258778	0.04791032
KLK7	kallikrein-related peptidase 7	5650	ILMN_174557	-0.7075403	0.00353032	-0.5757744	0.04084235
PLA2G16	phospholipase A2, group XVI	11145	ILMN_166771	-0.6960699	0.00281378	-0.6899545	0.0104823
ELOVL6	ELOVL fatty acid elongase 6	79071	ILMN_211118	-0.6727622	0.00792305	-0.6526218	0.03055515
LAMA3	laminin, alpha 3	3909	ILMN_170424	-0.670769	0.02031334	-0.8177877	0.0193704
CLIC4	chloride intracellular channel 4	25932	ILMN_167125	-0.6701399	0.00777512	-0.7301453	0.01522666
LAMA3	laminin, alpha 3	3909	ILMN_168889	-0.6510252	0.03134431	-1.0828059	0.00425672
SDC1	syndecan 1	6382	ILMN_181530	-0.6396137	0.00536134	-0.6052988	0.02592263
DSC2	desmocollin 2	1824	ILMN_166311	-0.6381088	0.00800621	-0.6020808	0.03440359
ACLY	ATP citrate lyase	47	ILMN_174901	-0.6203098	0.03100505	-0.7299685	0.03511605
AEN	apoptosis enhancing nuclease	64782	ILMN_181428	-0.6163406	0.01713013	-0.7029385	0.0235862
KRT6B	keratin 6B, type II	3854	ILMN_169877	-0.6126749	0.00766834	-0.7378259	0.0072612
KRT6C	keratin 6C, type II	286887	ILMN_175457	-0.6088326	0.00619183	-0.6104754	0.0202103
KRT16	keratin 16, type I	3868	ILMN_222816	-0.5940951	0.02299086	-1.169535	0.00083544
MSMO1	methylsterol monooxygenase 1	6307	ILMN_172088	-0.5680877	0.01733168	-0.9386443	0.00206309
ACTG2	actin, gamma 2, smooth muscle, enteric	72	ILMN_179532	-0.5631561	0.02053954	-0.6038613	0.03868784
SUSD2	sushi domain containing 2	56241	ILMN_169327	-0.5551115	0.03134431	-0.6456533	0.03833581
AMY1A	Amylase, Alpha 1A (Salivary)	276	ILMN_229476	-0.5510957	0.02112506	-0.5941947	0.03868784
RHPN2	rhophilin, Rho GTPase binding protein 2	85415	ILMN_175314	-0.5492825	0.01751972	-1.1020212	0.00048903
ALG13	ALG13, UDP-N-acetylglucosaminyltransferase subunit	79868	ILMN_176145	-0.5093844	0.03007361	-0.5804038	0.04102096
CAMK2N1	calcium/calmodulin-dependent protein kinase II inhibitor 1	55450	ILMN_179486	-0.5085462	0.02363425	-0.8827724	0.00214574
ACTR3	ARP3 actin-related protein 3 homolog (yeast)	10096	ILMN_165715	-0.497463	0.02244756	-0.5291919	0.04498074
SKA2	spindle and kinetochore associated complex subunit 2	348235	ILMN_324980	-0.4779017	0.0369661	-0.6051254	0.03024004
SNCA	synuclein, alpha (non A4 component of amyloid precursor)	6622	ILMN_170193	-0.4675343	0.04848173	-0.7837064	0.0069633
CRBN	cereblon	51185	ILMN_222573	0.44419284	0.04093207	0.61257768	0.0202103

Genes differentially expressed between MCF10A-scr and HOXA5 depleted cells that have an FDR < 0.05.

M6PR	mannose-6-phosphate receptor (cation dependent)	4074	ILMN_208435	0.44980885	0.04807991	0.55206608	0.04688306
RPL34	ribosomal protein L34	6164	ILMN_170687	0.45363558	0.03788979	0.87165005	0.00198803
ARHGDIB	Rho GDP dissociation inhibitor (GDI) beta	397	ILMN_167814	0.46531738	0.04138392	0.64605308	0.01998833
RELL1	RELT-like 1	768211	ILMN_323338	0.47538441	0.02854515	0.73666471	0.00569265
ANTXR1	anthrax toxin receptor 1	84168	ILMN_167037	0.48925564	0.04331091	0.64139985	0.0302158
EPB41L4A-A	EPB41L4A antisense RNA 1	114915	ILMN_324458	0.49679059	0.02715958	0.74371902	0.0067651
KIAA1644	KIAA1644	85352	ILMN_172888	0.49726378	0.02491949	0.60904232	0.02317837
PEMT	phosphatidylethanolamine N-methyltransferase	10400	ILMN_174580	0.49866621	0.02472493	0.58181184	0.03055515
CPVL	carboxypeptidase, vitellogenic-like	54504	ILMN_240075	0.50608576	0.02098935	0.81436189	0.00329615
LETMD1	LETM1 domain containing 1	25875	ILMN_175414	0.51540336	0.02299086	0.57890583	0.03364813
CYBRD1	cytochrome b reductase 1	79901	ILMN_208769	0.51555371	0.04127557	0.84697418	0.00636291
KLHDC3	kelch domain containing 3	116138	ILMN_173094	0.52900447	0.01803733	0.56754144	0.03365609
COL8A1	collagen, type VIII, alpha 1	1295	ILMN_240239	0.53344732	0.04771331	0.99046905	0.00366168
CPVL	carboxypeptidase, vitellogenic-like	54504	ILMN_168292	0.56388048	0.01107981	0.76038788	0.00498189
FBN2	fibrillin 2	2201	ILMN_167089	0.57894944	0.01165872	0.59362134	0.03055515
MUC1	mucin 1, cell surface associated	4582	ILMN_175699	0.61755276	0.00792305	1.0865568	0.00051381
DKK1	dickkopf WNT signaling pathway inhibitor 1	22943	ILMN_177333	0.63916825	0.00628866	0.59409254	0.03189396
MUC1	mucin 1, cell surface associated	4582	ILMN_237191	0.65099611	0.02761942	1.08913419	0.00360234
COL8A1	collagen, type VIII, alpha 1	1295	ILMN_168543	0.68238061	0.00296813	0.81992245	0.00324329
CBX6	chromobox homolog 6	23466	ILMN_169193	0.68665396	0.00772591	0.61619998	0.04456079
TNFRSF6B	Tumor Necrosis Factor Receptor Superfamily, Member 6b, De	8771	ILMN_233123	0.72029486	0.00354072	0.69398591	0.0162877
HSD17B10	hydroxysteroid (17-beta) dehydrogenase 10	3028	ILMN_169192	0.72638583	0.00240822	0.83203957	0.00360234
RPS29	ribosomal protein S29	6235	ILMN_229881	0.72854431	0.02098935	1.14478978	0.00370762
ANPEP	alanyl (membrane) aminopeptidase	290	ILMN_176383	0.75508367	0.01412231	0.85940153	0.0202103
PROS1	protein S (alpha)	5627	ILMN_167192	0.76124922	0.00193802	0.93676532	0.00175687
STMN3	stathmin-like 3	50861	ILMN_324411	0.77012297	0.0077385	0.78266328	0.0223109
SLC16A3	solute carrier family 16 (monocarboxylate transporter), membe	9123	ILMN_236402	0.77526681	0.00181746	0.66766743	0.01714984
TNFRSF6B	tumor necrosis factor receptor superfamily, member 6b, decoy	8771	ILMN_166182	0.78844692	0.00129551	0.62264836	0.02128
SLC26A2	solute carrier family 26 (anion exchanger), member 2	1836	ILMN_180149	0.85290469	0.00509148	0.72923569	0.04165771
DECR1	2,4-dienoyl CoA reductase 1, mitochondrial	1666	ILMN_172083	0.88858142	0.0006159	0.70778774	0.01078628
TNFRSF6B	tumor necrosis factor receptor superfamily, member 6b, decoy	8771	ILMN_233123	0.92965117	0.00059837	0.93478781	0.00214574

Genes differentially expressed between MCF10A-scr and HOXA5 depleted cells that have an FDR < 0.05.

MT1X	metallothionein 1X	4501	ILMN_177517	0.93918111	0.00152451	0.8704243	0.00905223
STEAP1	six transmembrane epithelial antigen of the prostate 1	26872	ILMN_173309	0.97028278	0.00030001	0.59263721	0.0329035
GJC2	gap junction protein, gamma 2, 47kDa	57165	ILMN_172304	1.05177703	0.00037414	0.95868156	0.00324329
ITGA10	integrin, alpha 10	8515	ILMN_170014	1.0630186	0.00014037	0.92669477	0.00197223
CYP4V2	cytochrome P450, family 4, subfamily V, polypeptide 2	285440	ILMN_205460	1.06843413	0.00061168	1.42649208	0.00024289
CD68	CD68 molecule	968	ILMN_171486	1.08671572	0.00044598	0.73124337	0.02549425
CREB3L2	cAMP responsive element binding protein 3-like 2	64764	ILMN_175109	1.19198339	0.00075991	0.83915567	0.02774872
HSD17B10	hydroxysteroid (17-beta) dehydrogenase 10	3028	ILMN_235597	1.24124483	0.00018478	1.15443748	0.00155767
MGC87042	NA		ILMN_166357	1.26190647	1.72E-05	0.82373688	0.00308692
DUSP23	dual specificity phosphatase 23	54935	ILMN_165946	1.35313787	9.68E-06	0.62562693	0.01881754
SRPX	sushi-repeat containing protein, X-linked	8406	ILMN_170948	1.42376485	5.16E-05	0.71930987	0.03189396
DYNLRB1	dynein, light chain, roadblock-type 1	83658	ILMN_176676	1.54878694	3.35E-06	1.59158712	5.41E-05
CYP1B1	cytochrome P450, family 1, subfamily B, polypeptide 1	1545	ILMN_169333	1.55916579	0.00017069	0.91445269	0.03015731
SCG5	secretogranin V (7B2 protein)	6447	ILMN_206577	1.56943716	3.51E-05	0.832108	0.0193704
DYNLRB1	dynein, light chain, roadblock-type 1	83658	ILMN_170356	1.59867027	1.97E-06	1.37306001	7.01E-05
AKR1C3	aldo-keto reductase family 1, member C3	8644	ILMN_171312	1.60590304	1.44E-05	0.64533716	0.04891569
DNER	delta/notch-like EGF repeat containing	92737	ILMN_179167	1.61320813	3.05E-05	0.8741671	0.01579879
GNG11	guanine nucleotide binding protein (G protein), gamma 11	2791	ILMN_178241	1.64916901	4.69E-06	0.76921219	0.0097746
DYNLRB1	dynein, light chain, roadblock-type 1	83658	ILMN_215743	1.66546724	1.84E-06	1.4040739	7.01E-05
DYNLRB1	dynein, light chain, roadblock-type 1	83658	ILMN_165805	1.71290812	1.97E-06	1.41105647	7.51E-05
AKR1C2	aldo-keto reductase family 1, member C2	1646	ILMN_168775	2.20364907	5.25E-06	0.7929915	0.04884242
AKR1C2	aldo-keto reductase family 1, member C2	1646	ILMN_241233	2.90562022	2.08E-08	1.08938662	0.00083544
DCN	decorin	1634	ILMN_176822	4.1434441	7.85E-10	0.69952483	0.01740298
DCN	decorin	1634	ILMN_234714	4.43815	7.85E-10	0.84830025	0.01011829

Bibliography

Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF (2003). Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* **100:** 3983-3988.

Alexander T, Nolte C, Krumlauf R (2009). Hox genes and segmentation of the hindbrain and axial skeleton. *Annu Rev Cell Dev Biol* **25:** 431-456.

Aubin J, Lemieux M, Tremblay M, Berard J, Jeannotte L (1997). Early postnatal lethality in Hoxa-5 mutant mice is attributable to respiratory tract defects. *Dev Biol* **192:** 432-445.

Berube-Simard FA, Jeannotte L (2014). Hoxa5/Cre transgenic mice: novel tools for regional deletion along the anterior-posterior axis. *Genesis* **52:** 149-156.

Chen H, Chung S, Sukumar S (2004). HOXA5-induced apoptosis in breast cancer cells is mediated by caspases 2 and 8. *Mol Cell Biol* **24:** 924-935.

Chen H, Zhang H, Lee J, Liang X, Wu X, Zhu T *et al* (2007). HOXA5 acts directly downstream of retinoic acid receptor beta and contributes to retinoic acid-induced apoptosis and growth inhibition. *Cancer Res* **67:** 8007-8013.

Chute JP, Muramoto GG, Whitesides J, Colvin M, Safi R, Chao NJ *et al* (2006). Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells. *Proc Natl Acad Sci U S A* **103:** 11707-11712.

Coradini D, Boracchi P, Oriana S, Biganzoli E, Ambrogi F (2014). Cell identity disruption in breast cancer precursors. *Anticancer Res* **34:** 1307-1319.

Debnath J, Muthuswamy SK, Brugge JS (2003). Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. *Methods* **30:** 256-268.

Desmedt C, Piette F, Loi S, Wang Y, Lallemand F, Haibe-Kains B *et al* (2007). Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. *Clin Cancer Res* **13:** 3207-3214.

Dontu G, Abdallah WM, Foley JM, Jackson KW, Clarke MF, Kawamura MJ *et al* (2003). In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev* **17:** 1253-1270.

Dubrulle J, McGrew MJ, Pourquie O (2001). FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell* **106:** 219-232.

Garin E, Lemieux M, Coulombe Y, Robinson GW, Jeannotte L (2006). Stromal Hoxa5 function controls the growth and differentiation of mammary alveolar epithelium. *Dev Dyn* **235:** 1858-1871.

Ghiaur G, Yegnasubramanian S, Perkins B, Gucwa JL, Gerber JM, Jones RJ (2013). Regulation of human hematopoietic stem cell self-renewal by the microenvironment's control of retinoic acid signaling. *Proc Natl Acad Sci U S A* **110:** 16121-16126.

Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M *et al* (2007). ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell* **1:** 555-567.

Gyorffy B, Lanczky A, Eklund AC, Denkert C, Budczies J, Li Q *et al* (2010). An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res Treat* **123:** 725-731.

Haberman AS, Isaac DD, Andrew DJ (2003). Specification of cell fates within the salivary gland primordium. *Dev Biol* **258:** 443-453.

Henderson GS, van Diest PJ, Burger H, Russo J, Raman V (2006). Expression pattern of a homeotic gene, HOXA5, in normal breast and in breast tumors. *Cell Oncol* **28:** 305-313.

Henderson KD, Andrew DJ (2000). Regulation and function of Scr, exd, and hth in the Drosophila salivary gland. *Dev Biol* **217:** 362-374.

Honeth G, Bendahl PO, Ringner M, Saal LH, Gruvberger-Saal SK, Lovgren K *et al* (2008). The CD44+/CD24- phenotype is enriched in basal-like breast tumors. *Breast Cancer Res* **10:** R53.

Jeanes A, Gottardi CJ, Yap AS (2008). Cadherins and cancer: how does cadherin dysfunction promote tumor progression? *Oncogene* **27:** 6920-6929.

Jin K, Kong X, Shah T, Penet MF, Wildes F, Sgroi DC *et al* (2012). The HOXB7 protein renders breast cancer cells resistant to tamoxifen through activation of the EGFR pathway. *Proc Natl Acad Sci U S A* **109:** 2736-2741.

Kay R, Rosten PM, Humphries RK (1991). CD24, a signal transducer modulating B cell activation responses, is a very short peptide with a glycosyl phosphatidylinositol membrane anchor. *J Immunol* **147:** 1412-1416.

Keller PJ, Lin AF, Arendt LM, Klebba I, Jones AD, Rudnick JA *et al* (2010). Mapping the cellular and molecular heterogeneity of normal and malignant breast tissues and cultured cell lines. *Breast Cancer Res* **12:** R87.

Kenny PA, Lee GY, Myers CA, Neve RM, Semeiks JR, Spellman PT *et al* (2007). The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression. *Mol Oncol* **1:** 84-96.

Klopp AH, Lacerda L, Gupta A, Debeb BG, Solley T, Li L *et al* (2010). Mesenchymal stem cells promote mammosphere formation and decrease E-cadherin in normal and malignant breast cells. *PLoS One* **5:** e12180.

Kominsky SL, Argani P, Korz D, Evron E, Raman V, Garrett E *et al* (2003). Loss of the tight junction protein claudin-7 correlates with histological grade in both ductal carcinoma in situ and invasive ductal carcinoma of the breast. *Oncogene* **22:** 2021-2033.

Kondo T, Takahashi N, Muramatsu M (1992). The regulation of the murine Hox-2.5 gene expression during cell differentiation. *Nucleic Acids Res* **20:** 5729-5735.

Korangath P, Teo WW, Sadik H, Han L, Mori N, Huijts CM *et al* (2015). Targeting Glutamine Metabolism in Breast Cancer with Aminooxyacetate. *Clin Cancer Res*.

Kramer F, White K, Kubbies M, Swisshelm K, Weber BH (2000). Genomic organization of claudin-1 and its assessment in hereditary and sporadic breast cancer. *Hum Genet* **107:** 249-256.

Krtolica A, Parrinello S, Lockett S, Desprez PY, Campisi J (2001). Senescent fibroblasts promote epithelial cell growth and tumorigenesis: a link between cancer and aging. *Proc Natl Acad Sci U S A* **98:** 12072-12077.

Kuperwasser C, Chavarria T, Wu M, Magrane G, Gray JW, Carey L *et al* (2004). Reconstruction of functionally normal and malignant human breast tissues in mice. *Proc Natl Acad Sci U S A* **101:** 4966-4971.

Ma XJ, Dahiya S, Richardson E, Erlander M, Sgroi DC (2009). Gene expression profiling of the tumor microenvironment during breast cancer progression. *Breast Cancer Res* **11:** R7.

Mani SA, Guo W, Liao MJ, Eaton EN, Ayyanan A, Zhou AY *et al* (2008). The epithelial-mesenchymal transition generates cells with properties of stem cells. *Cell* **133:** 704-715.

Marshall H, Morrison A, Studer M, Popperl H, Krumlauf R (1996). Retinoids and Hox genes. *Faseb J* **10:** 969-978.

Martin TA, Jiang WG (2009). Loss of tight junction barrier function and its role in cancer metastasis. *Biochim Biophys Acta* **1788:** 872-891.

Martin TA, Mansel RE, Jiang WG (2010). Loss of occludin leads to the progression of human breast cancer. *Int J Mol Med* **26:** 723-734.

May CD, Sphyris N, Evans KW, Werden SJ, Guo W, Mani SA (2011). Epithelial-mesenchymal transition and cancer stem cells: a dangerously dynamic duo in breast cancer progression. *Breast Cancer Res* **13:** 202.

Novak P, Jensen T, Oshiro MM, Wozniak RJ, Nouzova M, Watts GS *et al* (2006). Epigenetic inactivation of the HOXA gene cluster in breast cancer. *Cancer Res* **66:** 10664-10670.

Padua D, Zhang XH, Wang Q, Nadal C, Gerald WL, Gomis RR *et al* (2008). TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* **133:** 66-77.

Petkova N, Hennenlotter J, Sobiesiak M, Todenhofer T, Scharpf M, Stenzl A *et al* (2013). Surface CD24 distinguishes between low differentiated and transit-amplifying cells in the basal layer of human prostate. *Prostate* **73:** 1576-1590.

Ponti D, Costa A, Zaffaroni N, Pratesi G, Petrangolini G, Coradini D *et al* (2005). Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties. *Cancer Res* **65:** 5506-5511.

Pratt MA, Langston AW, Gudas LJ, McBurney MW (1993). Retinoic acid fails to induce expression of Hox genes in differentiation-defective murine embryonal carcinoma cells carrying a mutant gene for alpha retinoic acid receptor. *Differentiation* **53:** 105-113.

Qiao Y, Jiang X, Lee ST, Karuturi RK, Hooi SC, Yu Q (2011). FOXQ1 regulates epithelial-mesenchymal transition in human cancers. *Cancer Res* **71:** 3076-3086.

Radisky DC, LaBarge MA (2008). Epithelial-mesenchymal transition and the stem cell phenotype. *Cell Stem Cell* **2:** 511-512.

Raman V, Martensen SA, Reisman D, Evron E, Odenwald WF, Jaffee E *et al* (2000). Compromised HOXA5 function can limit p53 expression in human breast tumours. *Nature* **405:** 974-978.

Ribeiro AS, Paredes J (2014). P-Cadherin Linking Breast Cancer Stem Cells and Invasion: A Promising Marker to Identify an "Intermediate/Metastable" EMT State. *Front Oncol* **4:** 371.

Schwitalla S, Fingerle AA, Cammareri P, Nebelsiek T, Goktuna SI, Ziegler PK *et al* (2013). Intestinal tumorigenesis initiated by dedifferentiation and acquisition of stem-cell-like properties. *Cell* **152:** 25-38.

Shah N, Sukumar S (2010). The Hox genes and their roles in oncogenesis. *Nat Rev Cancer* **10:** 361-371.

Sheridan C, Kishimoto H, Fuchs RK, Mehrotra S, Bhat-Nakshatri P, Turner CH *et al* (2006). CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis. *Breast Cancer Res* **8:** R59.

Simeone A, Acampora D, Arcioni L, Andrews PW, Boncinelli E, Mavilio F (1990). Sequential activation of HOX2 homeobox genes by retinoic acid in human embryonal carcinoma cells. *Nature* **346:** 763-766.

Sleeman KE, Kendrick H, Ashworth A, Isacke CM, Smalley MJ (2006). CD24 staining of mouse mammary gland cells defines luminal epithelial, myoepithelial/basal and non-epithelial cells. *Breast Cancer Res* **8:** R7.

Stornaiuolo A, Acampora D, Pannese M, D'Esposito M, Morelli F, Migliaccio E *et al* (1990). Human HOX genes are differentially activated by retinoic acid in embryonal

carcinoma cells according to their position within the four loci. *Cell Differ Dev* **31:** 119-127.

Strickland S, Mahdavi V (1978). The induction of differentiation in teratocarcinoma stem cells by retinoic acid. *Cell* **15:** 393-403.

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA *et al* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102:** 15545-15550.

Swisshelm K, Machl A, Planitzer S, Robertson R, Kubbies M, Hosier S (1999). SEMP1, a senescence-associated cDNA isolated from human mammary epithelial cells, is a member of an epithelial membrane protein superfamily. *Gene* **226:** 285-295.

Taylor HS, Vanden Heuvel GB, Igarashi P (1997). A conserved Hox axis in the mouse and human female reproductive system: late establishment and persistent adult expression of the Hoxa cluster genes. *Biol Reprod* **57:** 1338-1345.

Tokes AM, Kulka J, Paku S, Szik A, Paska C, Novak PK *et al* (2005). Claudin-1, -3 and -4 proteins and mRNA expression in benign and malignant breast lesions: a research study. *Breast Cancer Res* **7:** R296-305.

van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M *et al* (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415:** 530-536.

Wang CC, Su KY, Chen HY, Chang SY, Shen CF, Hsieh CH *et al* (2015). HOXA5 Inhibits Metastasis via Regulating Cytoskeletal Remodelling and Associates with Prolonged Survival in Non-Small-Cell Lung Carcinoma. *PLoS One* **10:** e0124191.

Wang H, Meyer CA, Fei T, Wang G, Zhang F, Liu XS (2013). A systematic approach identifies FOXA1 as a key factor in the loss of epithelial traits during the epithelial-to-mesenchymal transition in lung cancer. *BMC Genomics* **14:** 680.

Wang YA, Shen K, Wang Y, Brooks SC (2005). Retinoic acid signaling is required for proper morphogenesis of mammary gland. *Dev Dyn* **234:** 892-899.

Watson RE, Curtin GM, Hellmann GM, Doolittle DJ, Goodman JI (2004). Increased DNA methylation in the HoxA5 promoter region correlates with decreased expression of the gene during tumor promotion. *Mol Carcinog* **41:** 54-66.

Yan W, Chen Y, Yao Y, Zhang H, Wang T (2013). Increased invasion and tumorigenicity capacity of CD44+/CD24- breast cancer MCF7 cells in vitro and in nude mice. *Cancer Cell Int* **13:** 62.

Zhang H, Meng F, Liu G, Zhang B, Zhu J, Wu F *et al* (2011). Forkhead transcription factor foxq1 promotes epithelial-mesenchymal transition and breast cancer metastasis. *Cancer Res* **71:** 1292-1301.

Zwartkruis F, Kruyt F, van der Saag PT, Meijlink F (1993). Induction of HOX-2 genes in P19 embryocarcinoma cells is dependent on retinoic acid receptor alpha. *Exp Cell Res* **205:** 422-425.

Curriculum Vitae

Wei Wen Teo

Nationality: Malaysian

D.O.B: 24 April 1980

Email: wteo2@jhmi.edu

weein24@gmail.com

LANGUAGES

- English
- Chinese
- Malay

EDUCATION AND EXPERIENCE

- PhD Candidate — current
  Pathobiology, Johns Hopkins School of Medicine

- Research Specialist — 2009
  Breast cancer Program SKCCC, Johns Hopkins

- Research Assistant — 2007
  Johns Hopkins Hospital Singapore

- Laboratory Officer — 2006
  Oncology Research Institute, National University Singapore

- Laboratory Officer — 2005

Gastroenterology Department, National University Singapore

- Biochemistry, bachelor degree 2003
  National University Singapore

ACCOMPLISHMENTS

- Recipient of Margaret Lee international fellowship from Pathobiology graduate Program at Johns Hopkins School of Medicine.
- Co-patentee of cMethDNA technology at Breast Cancer Program of Johns Hopkins University.
- 2015 Pathology Young Investigators' Day Award, Excellence in Basic Research.
- 2015 Johns Hopkins SKCCC Fellow Research Day Award, Honorable Mention for Basic Research.
- Poster presenter at American Association for Cancer Research Annual Meeting 2015.
- Poster presenter at Annual Safeway Breast Cancer Retreat from 2012 to 2015.
- Poster presenter at Annual Pathobiology Retreat from 2011 to 2014 and student speaker in 2014.

PEER-REVIEWED PUBLICATIONS

1. Wu JM, Fackler MJ, Halushka MK, Molavi DW, Taylor ME, **Teo WW**, Griffin C, Fetting J, Davidson NE, De Marzo AM, Hicks JL, Chitale D, Ladanyi M, Sukumar S, Argani P. Heterogeneity of breast cancer metastases: comparison of therapeutic target expression and promoter methylation between primary tumors and their multifocal metastases. Clin Cancer Res. 2008 Apr 1; 14(7):1938-46.

2. Lee JS, Fackler MJ, **Teo WW**, Lee JH, Choi C, Park MH, Yoon JH, Zhang Z, Argani P, Sukumar S. Quantitative promoter hypermethylation profiles of ductal carcinoma in situ in North American and Korean women: Potential applications for diagnosis. Cancer Biol Ther. 2008 Sep; 7(9):1398-406.

3. Fackler MJ, Rivers A, **Teo WW**, Mangat A, Taylor E, Zhang Z, Goodman S, Argani P, Nayar R, Susnik B, Sukumar S, Khan SA. Hypermethylated genes as biomarkers of cancer in women with pathologic nipple discharge. Clin Cancer Res. 2009 Jun 1; 15(11):3802-11.

4. Lo PK, Watanabe H, Cheng PC, **Teo WW**, Liang X, Argani P, Lee JS, Sukumar S.MethySYBR, a novel quantitative PCR assay for the dual analysis of DNA methylation and CpG methylation density. J Mol Diagn. 2009 Sep; 11(5):400-14.

5. Fackler MJ, Umbricht CB, Williams D, Argani P, Cruz LA, Merino VF, **Teo WW**, Zhang Z, Huang P, Visvananthan K, Marks J, Ethier S, Gray JW, Wolff AC, Cope LM, Sukumar S. Genome-wide methylation analysis identifies genes specific to breast cancer hormone receptor status and risk of recurrence. 2011 Oct 1; 71(19):6195-207.

6. Fackler MJ, Lopez Bujanda Z, Umbricht C, **Teo WW**, Cho S, Zhang Z, Visvanathan K, Jeter S, Argani P, Wang C, Lyman JP, de Brot M, Ingle JN, Boughey J, McGuire K, King TA, Carey LA, Cope L, Wolff AC, Sukumar S. Novel methylated biomarkers and a robust assay to detect circulating tumor DNA in metastatic breast cancer. Cancer Res. 2014 Apr 15; 74(8):2160-70.

7. **Teo WW** and Sukumar S. Combining the strength of genomics, nanoparticle technology, and direct intraductal delivery for breast cancer treatment. (Viewpoint) *Breast Cancer Research* 2014, **16**:306.

8. Han L, Diehl A, Nguyen N, Korangath P, **Teo WW**, Cho S, Kominsky S, Huso D, Feigenbaum L, Rein A, Argani P, Landberg G, Gessler M, Sukumar S. The Notch pathway inhibits TGF-β signaling in breast cancer through HEYL-mediated crosstalk. Cancer Res. 2014 Nov 15;74(22):6509-18

9. Korangath P, **Teo WW**, Sadik H, Han L, Mori N, Huijts CM, Wildes F, Bharti S, Zhang Z, Santa-Maria CA, Tsai H, Dang CV, Stearns V, Bhujwalla ZM, Sukumar S. Targeting Glutamine Metabolism in Triple Negative Breast Cancer with Aminooxyacetate. Clin Cancer Res. 2015 Mar 26

10. Jin K, Park S, **Teo WW**, Korangath P, Cho S, Yoshida T, Györffy B, Goswami C, Nakshatri H, Cruz L, Zhou W, Ji H, Su Y, Ekram M, Polyak K, and Sukumar S. HOXB7 is an ERα cofactor in the activation of HER2 and multiple ER target genes leading to endocrine resistance. Cancer Discov. 2015 Sep;5(9):944-59.

11. **Teo WW**, Merino V, Cho S, Korangath P, Sukumar S. HOXA5, a Cell State Transition Regulator, Inhibits Tumor Initiation and Cancer Progression. (Manuscript submitted).